

MULTICELLULAR SYSTEMS BIOLOGY OF DEVELOPMENT

DISSERTATION

zur Erlangung des akademischen Grades

Doctor of Philosophy  
(Ph.D.)

vorgelegt

der Fakultät Mathematik und Naturwissenschaften  
der Technischen Universität Dresden

von

Walter de Back

geboren am 28.09.1978  
in Terneuzen, die Niederlande

eingereicht am 09.07.2015  
verteidigt am 03.11.2015

Gutachter:  
Prof. em. Dr. Hans Meinhardt  
Prof. Dr. Christian Dahmann

Die Dissertation wurde am  
Zentrum für Informationsdienste und Hochleistungsrechnen (ZIH) angefertigt.



---

## SUMMARY

---

Embryonic development depends on the precise coordination of cell fate specification, patterning and morphogenesis. Although great strides have been made in the molecular understanding of each of these processes, how their interplay governs the formation of complex tissues remains poorly understood. New techniques for experimental manipulation and image quantification enable the study of development in unprecedented detail, resulting in new hypotheses on the interactions between known components. By expressing these hypotheses in terms of rules and equations, computational modeling and simulation allows one to test their consistency against experimental data. However, new computational methods are required to represent and integrate the network of interactions between gene regulation, signaling and biomechanics that extend over the molecular, cellular and tissue scales.

In this thesis, I present a framework that facilitates computational modeling of multi-scale multicellular systems and apply it to investigate pancreatic development and the formation of vascular networks. This framework is based on the integration of discrete cell-based models with continuous models for intracellular regulation and intercellular signaling. Specifically, gene regulatory networks are represented by differential equations to analyze cell fate regulation; interactions and distributions of signaling molecules are modeled by reaction-diffusion systems to study pattern formation; and cell-cell interactions are represented in cell-based models to investigate morphogenetic processes. A cell-centered approach is adopted that facilitates the integration of processes across the scales and simultaneously constrains model complexity.

The computational methods that are required for this modeling framework have been implemented in the software platform Morpheus. This modeling and simulation environment enables the development, execution and analysis of multi-scale models of multicellular systems. These models are represented in a new domain-specific markup language that separates the biological model from the computational methods and facilitates model storage and exchange. Together with a user-friendly graphical interface, Morpheus enables computational modeling of complex developmental processes without programming and thereby widens its accessibility for biologists.

To demonstrate the applicability of the framework to problems in developmental biology, two case studies are presented that address different aspects of the interplay between cell fate specification, patterning and morphogenesis. In the first, I focus on the interplay between cell fate stability and intercellular signaling. Specifically, two studies

are presented that investigate how mechanisms of cell-cell communication affect cell fate regulation and spatial patterning in the pancreatic epithelium. Using bifurcation analysis and simulations of spatially coupled differential equations, it is shown that intercellular communication results in a multistability of gene expression states that can explain the scattered spatial distribution and low cell type ratio of nascent islet cells. Moreover, model analysis shows that disruption of intercellular communication induces a transition between gene expression states that can explain observations of *in vitro* transdifferentiation from adult acinar cells into new islet cells. These results emphasize the role of the multicellular context in cell fate regulation during development and may be used to optimize protocols for cellular reprogramming.

The second case study focuses on the feedback between patterning and morphogenesis in the context of the formation of vascular networks. Integrating a cell-based model of endothelial chemotaxis with a reaction-diffusion model representing signaling molecules and extracellular matrix, it is shown that vascular network patterns with realistic morphometry can arise when signaling factors are retained by cell-modified matrix molecules. Through the validation of this model using *in vitro* assays, quantitative estimates are obtained for kinetic parameters that, when used in quantitative model simulations, confirm the formation of vascular networks under measured biophysical conditions. These results demonstrate the key role of the extracellular matrix in providing spatial guidance cues, a fact that may be exploited to enhance vascularization of engineered tissues.

Together, the modeling framework, software platform and case studies presented in this thesis demonstrate how cell-centered computational modeling of multi-scale and multicellular systems provide powerful tools to help disentangle the complex interplay between cell fate specification, patterning and morphogenesis during embryonic development.



---

## PUBLICATIONS

---

- A. Köhn-Luque, W. de Back, J. Starruß, A. Mattioli, A. Deutsch, J-M. Pérez-Pomares and M. A. Herrero. Early embryonic vascular patterning by matrix-mediated paracrine signalling: a mathematical model study. *PLoS ONE* 6(9):e24175, 2011.
- A. Voß-Böhme, J. Starruß and W. de Back. Cellular Potts model. In: *Encyclopedia of systems biology*, W. Dubitzky, O. Wolkenhauer, H. Yokota, K.-H. Cho (eds.). Springer, 2012.
- W. de Back, J. X. Zhou and L. Brusch. On the role of lateral stabilization during early patterning in the pancreas. *Journal of the Royal Society Interface*, 10(79): 20120766, 2012.
- W. de Back, R. Zimm and L. Brusch. Transdifferentiation of pancreatic cells by loss of contact-mediated signaling. *BMC Systems Biology* 7:77, 2013.
- A. Köhn-Luque, W. de Back, Y. Yamaguchi, K. Yoshimura, M. A. Herrero and T. Miura. Dynamics of VEGF matrix-retention in vascular network patterning. *Physical Biology*, 10:066007, 2013.
- J. Starruß\*, W. de Back\*, L. Brusch and A. Deutsch. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics*, 30(9):1331-1332, 2014. (\*equal contribution)
- M. Herberg, T. Zerjatke, W. de Back, I. Glauche and I. Roeder. Image-based quantification and mathematical modeling of spatial heterogeneity in mESC colonies. *Cytometry Part A*, 10.1002/cyto.a.22598, 2015.
- W. de Back, A. Deutsch, D. Drasdo, A. Funahashi, A. M. Uhrmacher. Towards a standard exchange format for spatial, multilevel multicellular models. In: *Dagstuhl Reports: Multiscale Spatial Computational Systems Biology*, D. Gilbert, M. Heiner, K. Takahashi, A. M. Uhrmacher (eds.). 4(11):214-225, 2015.



---

## CONTENTS

---

1	INTRODUCTION	1
1.1	Systems biology . . . . .	5
1.2	Multicellular systems biology of development . . . . .	6
1.3	Overview of thesis . . . . .	9
1.4	Supplementary online material . . . . .	11
I	COMPUTATIONAL METHODS	13
2	INTRODUCTION	14
2.1	Approach . . . . .	15
2.2	Computational methods . . . . .	17
2.3	Software environments . . . . .	22
3	MORPHEUS: A MODELING ENVIRONMENT FOR MULTISCALE AND MULTICELLULAR SYSTEMS BIOLOGY	25
3.1	Introduction . . . . .	25
3.2	Modeling formalisms . . . . .	26
3.3	Graphical user interface . . . . .	36
3.4	MorpheusML . . . . .	39
3.5	Automated model integration . . . . .	44
3.6	Software architecture and extensibility . . . . .	47
3.7	Conclusions . . . . .	50
II	CELL FATE REGULATION AND PATTERNING	53
4	INTRODUCTION	54
4.1	Cellular plasticity and reprogramming . . . . .	54
4.2	Mathematical modeling of transcriptional regulation . . . . .	56
4.3	Mathematical modeling of intercellular communication . . . . .	58
4.4	Cell fates, patterns and reprogramming in the pancreas . . . . .	60
4.5	Conclusions . . . . .	60
5	ON THE ROLE OF LATERAL STABILIZATION DURING EARLY PATTERNING IN THE PANCREAS	62
5.1	Introduction . . . . .	62

## CONTENTS

5.2	Methods . . . . .	64
5.3	Results and Discussion . . . . .	68
5.4	Conclusion . . . . .	78
6	TRANSDIFFERENTIATION OF PANCREATIC CELLS BY LOSS OF CONTACT-MEDIATED SIGNALING	80
6.1	Introduction . . . . .	80
6.2	Methods . . . . .	82
6.3	Results . . . . .	85
6.4	Discussion and Conclusion . . . . .	92
III	MORPHOGENESIS	95
7	INTRODUCTION	96
7.1	The vascular system . . . . .	96
7.2	Vasculogenesis . . . . .	97
7.3	Mathematical models of vascular network formation . . . . .	98
7.4	Paracrine model of vascular network formation . . . . .	101
7.5	Conclusions . . . . .	103
8	EARLY EMBRYONIC VASCULAR PATTERNING BY MATRIX-MEDIATED PARACRINE SIGNALLING	105
8.1	Introduction . . . . .	105
8.2	Methods . . . . .	108
8.3	Results . . . . .	115
8.4	Discussion . . . . .	124
9	DYNAMICS OF VEGF MATRIX-RETENTION IN VASCULAR NETWORK PATTERNING	128
9.1	Introduction . . . . .	128
9.2	Materials and Methods . . . . .	130
9.3	Results and Discussion . . . . .	139
9.4	Conclusions . . . . .	145
10	CONCLUSIONS	149
A	APPENDICES	159
A.1	Computational performance and scalability . . . . .	159
A.2	Observation limit for a FRAP experiment . . . . .	161

## CONTENTS

A.3 Derivation of the FRAP recovery curve . . . . .	165
BIBLIOGRAPHY	167





---

## INTRODUCTION

---

The adult human body consists of an estimated  $3.72 \cdot 10^{13}$  cells<sup>1</sup> that can be grouped in about 200 specialized cell types<sup>2</sup> forming intricate tissue architectures that make up 60 organs<sup>3</sup>. How such a complex multicellular organism can develop from a single fertilized egg cell has fascinated developmental biologists over centuries. With the rise of molecular biology, many of the molecular components and mechanisms have been revealed by which embryos are patterned, cells to acquire different fates and tissues adopt their defined shapes. However, how these mechanisms interact with each other to ensure the robust self-organization of tissues is still poorly understood. In addition to careful experimentation and quantification, mathematical and computational modeling is increasingly important to formulate assumptions and test hypotheses. However, a lack of suitable methods and tools to represent multiscale multicellular systems hampers these studies. Here, a framework called ‘multicellular systems biology’ is proposed to explore the effects of the interplay between genetic, chemical and mechanical interactions on tissue development.

Over the last decades, molecular approaches in developmental biology, from developmental genetics to tissue morphogenesis, have uncovered a plethora of molecular components that regulate embryonic development. This has led to the discovery of a wide variety of regulatory mechanisms and the appreciation of the fact that developmental processes are regulated at many levels. At the same time, it has become clear that surprisingly few, well-conserved, pathways are involved. Intercellular signaling involves only a handful of families of signaling factors using mostly linear pathways<sup>4</sup>; morphogenetic movements are predominantly regulated by a single family of adhesion molecules<sup>5,6</sup>; and tissue-specific gene expression is controlled by a small set of master regulators.

Interestingly, the misexpression of a single regulator gene or signaling factor has been shown to induce the ectopic formation of complex structures such as eyes on wings<sup>7</sup> or additional limbs<sup>8</sup>, showing that tissue formation can result from simple

local triggers. This is demonstrated even more elegantly in the *ex vivo* development of organ-like structures such as intestinal crypts<sup>9</sup> and optic cups<sup>10</sup> from stem cell cultures, in which external influences can be excluded. These results indicate that the formation of complex tissues is, for an important part, driven by self-organization<sup>11</sup>.

Therefore, the question arises how the limited number of key developmental mechanisms are combined in space and time to robustly generate the diversity of cell types and the variety of tissue architectures. Traditionally, development is seen as a sequence of more or less independent processes where patterned signals specify cell fates and the specified cells subsequently rearrange to acquire specific tissue shapes. Increasingly, however, it is realized that these are concurrent processes that occur simultaneously and on similar timescales, opening the possibility of regulatory feedbacks between cell fate specification, patterning and morphogenesis that go beyond the traditionally assumed sequence of causality<sup>12–14</sup>.

In gastrulation, for instance, many of the key signaling pathways are now known to simultaneously affect both cell fates and cell movement<sup>13</sup>. Therefore, cell movements may affect cell fate decisions by rearranging cells in order to expose them to subsequent sets of signals<sup>13</sup>. In stem cell biology, it is known that mechanical forces can control self-renewal and lineage specification<sup>15</sup>. The size of stem cell aggregates has been shown to control the induction of spontaneous symmetry breaking and self-organized axis formation<sup>16</sup>. In plant development, mechanical stresses generated by tissue growth were shown to affect the transport of gene-regulating growth hormones<sup>17</sup>. These are only few examples of the accumulating evidence that important feedbacks exist between morphogenesis, patterning and cell fate decisions. This imposes a new logic that could dramatically change our understanding of the regulation and robustness of developmental systems and change our interpretation of mutant phenotypes<sup>12</sup>. However, the biological mechanisms behind such interplay remain poorly understood.

On the one hand, new techniques are required to manipulate and visualize these highly dynamic developmental processes. Upcoming technologies such as targeted and conditional gene expression<sup>18,19</sup>, high-throughput genetic screens<sup>20</sup>, time-lapse imaging of living embryos<sup>21</sup> and quantitative image analysis<sup>22</sup> are crucial to provide high resolution quantitative data on these dynamics across spatiotemporal scales.

On the other hand, new methods for mathematical and computational modeling are needed to explore the consequences of such feedbacks since, generally, the behavior of complex dynamic systems goes far beyond human intuition. The large amounts of quantitative data that are being acquired are prompting scientists to new hypotheses explaining complex biological processes. Expressing these hypotheses in terms of rules and equations that can be simulated by a computer allows the exploration of non-intuitive consequences of complex interactions and feedback loops. Quantitative comparison of simulation results with data enables one to narrow down the set of possible hypotheses to a few plausible ones that can be tested experimentally (figure 1.1).



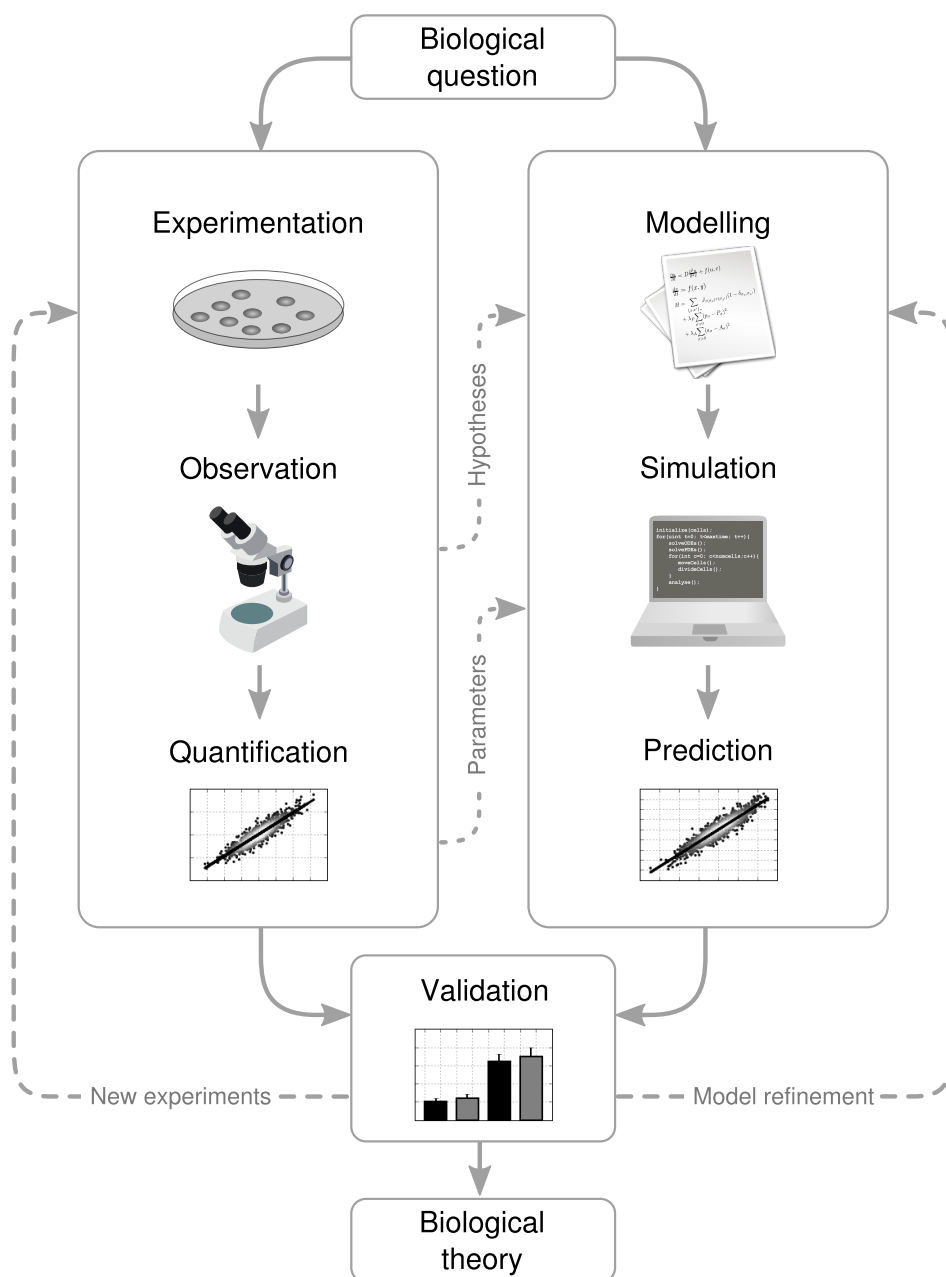


Figure 1.1.: **The role of computational modeling in quantitative biology.**  
(Caption on next page)

Figure 1.2.: To answer specific biological questions, observations and quantification of wet lab experiments are accompanied by the construction of mathematical and computational models that formally express hypotheses and assumptions. These models are based on first principles and/or on hypotheses derived from new experimental observations. When quantification of experimental data is used for model parameterization, simulations generate predictions that can be directly used for statistical validation against experimental data. Validation may enforce model refinement or suggest the acquisition of new experimental data. Through iteration between experimentation, quantification and modeling, increasing confidence in the validity of model assumptions eventually leads to the establishment of new biological theory.

To mathematically or computationally represent feedbacks between cell fate specification, patterning and morphogenesis, we can take advantage of a variety of existing modeling methods that are commonly used to represent distinct developmental mechanisms. Dynamic models, typically ordinary differential equations, of genetic regulatory networks allow us to understand the cell fate decisions of multipotent cells and how their trajectories through gene expression space is constrained by interactions between genes<sup>23</sup>. Spatial models such as reaction-diffusion models enable us to reveal how morphogen gradients influences these decisions and how the patterned gradients can spontaneously arise through interaction between diffusible morphogens<sup>24</sup>. And cell-based models, in which cells are represented as discrete and motile agents, allow us to grasp how specific tissue shapes emerge from the mechanical interactions between cells<sup>25</sup>.

In this thesis, a framework is presented that facilitates the exploration of such interplay. Conceptually, this framework is based on the integration of aforementioned methods to address specific questions on the effects of interactions and feedbacks between cell fate decisions, signaling and morphogenesis. Practically, a software platform has been developed that implements these computational methods and enables the flexible and user-friendly construction and simulation of multi-scale models of multicellular systems (part I). Two case studies are presented that illustrate the applicability of the approach to different types of problems in developmental biology. In the first, the crosstalk between cell fate specification and tissue patterning is studied in the context of development of the pancreas (part II). In the second case study, the interplay between patterning and morphogenesis is studied in the formation of vascular capillary networks (part III).

In the remainder of this introduction, I describe the background of the approach in which the perspective from systems biology is adopted in order to highlight the focus on dynamic interaction, integration and quantification. It is worth noting that similar concepts have recently been proposed in the context of dynamic imaging of animal development<sup>26</sup>, as ‘computational morphodynamics’ in the context of plant development<sup>27,28</sup> and as ‘cytosystems dynamics’ in the context of studies using organoids<sup>11</sup>.

## 1.1 SYSTEMS BIOLOGY

Biological organisms are entities that maintain their existence through the mutual interaction of its parts<sup>29</sup>, i.e. they are *systems*<sup>30</sup>. This view has become the driving force for an approach to the study of complex biological systems, called systems biology. Whereas the systems biological approach has had major influence on the fields of molecular biology and physiology, its impact on developmental biology has so far been much less pronounced. However, giving the current focus on quantification and dynamic modeling in understanding developing systems, the role of the systems biology is becoming increasingly important to developmental biology.

Systems biology is typically defined in contrast to the long dominant approach of reductionism in biological research. Whereas reductionist approaches have led to the successful identification and characterization of a large number of biological components, systems biology focuses on the integration of these components by looking at the dynamics of their interaction. Systems biology heavily relies on the mathematics of dynamical systems theory as well as computational simulations in order to understand the dynamics and emergent properties that arise through nonlinear interactions<sup>31</sup>. In addition, it emphasizes detailed quantification of experimental data in order to establish quantitative computational models that can provide testable predictions and reliably validate these models against experimental data.

Systems biology is set of common principles rather than a discipline<sup>32</sup> and has irremovable traces of its origins in two different fields, molecular cell biology and physiology, each of which can contribute to a multicellular systems biology of development.

### 1.1.1 Molecular systems biology

Through the fast developments in molecular biology and the rapid scale-up and automation of experimental techniques in the various 'omics', data on various levels of molecular cell biology, from genes and proteins to metabolites, has rapidly accumulated. To analyze these data, computational biological approaches, in particular bioinformatics and systems biology, have quickly been adopted. Bioinformatics provides methods for data management, storage and analysis to extract information about e.g. molecular interactions. Complementary to this, systems biology provides methods to gain insight into the dynamics of networks of such interactions. It analyzes the static topology of biochemical networks in order to identify control mechanisms such as positive and negative feedback loops. In addition, it focuses on the analysis of the dynamical behavior that arises from interactions within large-scale biochemical networks. As such, systems biology views the cell as a network of interacting molecular components. This has been influential in replacing the classical thinking in terms of pathways with the concepts imported from network science and dynamical systems theory<sup>33-36</sup>. However, given the focus on temporal dynamics of large molecular systems, spatial

dynamics and interactions with other spatiotemporal scales have largely been ignored in molecular systems biology.

### 1.1.2 Physiological systems biology

This is not true for the other branch of systems biology, originating from physiology. Here, the focus is on the regulation of biological functions and homeostasis through the interplay between levels of organization in the human body. Under the assumption that there is no privileged level of causation in biological systems, their regulation can only be understood as an interaction between these levels<sup>37</sup>. Therefore, rather than integration at the molecular scale, it concentrates on the integration of biological processes over temporal as well as spatial scales to understand organ or even whole-body physiology<sup>29,38,39</sup>. As a prominent example, the physiology of the heart has been modeled by coupling of the activity of ion channels to cell membranes potentials and to the spatial propagation of spiral waves over the ventricular tissue.

In order to integrate biological processes and computational models of a wide variety of scales, a pragmatic 'middle-out' approach is typically employed. In this approach, a certain level of biological organization is initially selected based on availability of data and represented in mathematical or computational terms. Additional processes at lower and higher levels are subsequently included as deemed relevant to the biological function under investigation<sup>40</sup>.

The discrepancy between these two branches of systems biology is perhaps best illustrated by their longterm goals. Molecular systems biology attempts to establish quantitative computational models that comprehensive describe the dynamics within the cell. This is exemplified in the recent establishment of a computational 'whole-cell' model of the life cycle of the bacterium *Mycoplasma genitalium* that, according to the authors, include all of its molecular components and their interactions<sup>41</sup>. In contrast, physiological systems biology aims to establish integrated computer models of the mechanical, physical and biochemical functions of a living human body. The goal of large-scale projects such as the Physiome project and the Virtual Physiological Human is to develop a computational framework for the quantitative description of biological processes in living systems across all relevant levels of structural and functional integration, from molecule to organism<sup>29</sup>.

## 1.2 MULTICELLULAR SYSTEMS BIOLOGY OF DEVELOPMENT

The differences between the molecular and physiological systems biology may provide a reason why systems biology has had less impact on developmental biology. Developmental biology attempts to understand how the combination of genetic, biochemical and mechanical interactions result in the patterning and shaping of embryonic tissues.

These processes cannot be studied by focusing purely on the molecular interactions at a single-cell level, nor do they require the inclusion of physiological processes at the level of whole organs or organisms. Instead, the most relevant processes for the study of developmental biology occur at the intermediate histological level. It is at this level or organization that cells mediate the exchange of information from biochemical and genetic regulation with intercellular signaling and tissue mechanics.

The branch of systems biology that focuses on the tissue level can be called 'multicellular systems biology'. Concentrating on the level between molecular cell biology and organismal physiology, it can adopt concepts and techniques from both the molecular and physiological branches of systems biology. On the one hand, it forms a natural and necessary extension of molecular-oriented systems biology towards the inclusion of spatial context and intercellular interactions. On the other hand, it adopts a multi-scale and spatial modeling approach from physiology-oriented systems biology.

However, unlike modeling in physiology, multicellular systems biology can restrict itself by using the cell as the central point of integration and the starting point of the middle-out approach. This convention has several advantages: (1) Compared to the vast variety of functions at the genetic and protein level, the number of behaviors that cells perform is more manageable and tractable. This allows the complex regulatory processes within tissues to be reduced, as a first approximation, to a small set of cellular behaviors, i.e. growth, division, polarity, adhesion, shape changes, signaling, differentiation, migration and death. (2) In line with the biological role of the cell as central regulator, the cell-based approach enables the coupling of cellular behaviors to processes at intracellular and intercellular levels. That is, it allows the study of the development and functioning of tissues by modeling the interactions of cells with each other, with their internal regulatory biochemical networks and with the extracellular microenvironment. (3) By establishing on a particular level of abstraction and associated modeling approaches, it enhances the development, comparison and standardization of mathematical and computational methodologies and thereby enhances scientific communication.

The use of mathematical models in developmental biology is, of course, not new. For instance, Alan Turing's seminal work on pattern formation by interacting morphogens, which remains an active topic in developmental biology to this day<sup>42,43</sup>, was published in 1952<sup>44</sup>. It can even be traced even back to 1917 with the publication of the book *On Growth and Form* in which D'Arcy Wentworth Thompson emphasized the role of mechanics and physics to explain morphogenesis<sup>45</sup>. A less well-known early example is the robotic model developed in 1912 by John Hammond Jr. and Benjamin Miessner whose control mechanism was explicitly based on the concept of heliotropism as proposed by the embryologist Jacques Loeb. Later, in a book published in 1918, Loeb took the heliotropic robot as providing important support for his mechanistic notions on phototactic plant development as well as animal behavior<sup>46</sup>.

Even though the use of models in developmental biology is about a century old, still, only a small fraction of research work done in the field uses mathematical and computational modeling. Why is this expected to change now? To answer this, it is instructive to draw a parallel with molecular systems biology. Although the important mathematical models of molecular kinetics were developed by Michaelis and Menten already in 1913<sup>47</sup>, systems biology only became a major approach 90 years later. Molecular systems biology emerged at a time when, apart from the mathematical tools, two other constraints were satisfied: (1) High-throughput technologies had been developed to a state that large amounts of quantitative data became available. (2) Computational methods and tools were sufficiently sophisticated that these data could be integrated in computational analysis and simulation. In developmental biology, these constraints are now also being met.

The recent developments in upcoming technologies such as gene editing<sup>19</sup>, single-cell transcriptomics<sup>48</sup>, 4D imaging of living specimen<sup>21</sup> and automated quantitative image analysis<sup>22</sup> are providing a wealth of quantitative data on the developmental regulation, at high spatial and temporal resolution. Furthermore, computational methods are now available to cope with these types of quantitative data, over different spatiotemporal scales. Whereas mathematical biology has long been dominated by first-principle studies, models can now be parameterized with measured biophysical quantities such as reaction kinetics, diffusion coefficients, binding rates, geometrical constraints from images, biophysical parameters of cell such as adhesion forces and compressibility constants, as well as behavioral parameters such as cell cycle lengths, migration speeds, apoptotic rates, etc. Based on such multiscale quantitative data, computational models can generate predictions that, when experimentally validated, can provide powerful evidence to demonstrate the consistency of a biological theory.

The dependency between quantitative data and computational modeling is bidirectional. On the one hand, the interpretation of quantitative multiscale data requires the construction of computational models to be able to link the activity of regulatory molecules to population- or tissue-level phenomena<sup>26</sup>. On the other hand, the explanatory potential of mechanistic computational models is related to the level in which they are able to provide quantitative predictions, allowing direct comparison to data obtained experimentally.

Nevertheless, theoretical modeling studies that are based on first principles rather than quantitative data remain important. In this respect, the novelty of multicellular systems biology are computational methods to couple theoretical models of cell fate specification, pattern formation and morphogenesis. These methods allow the exploration of consequences of hypothesized feedbacks between these processes, even before quantitative data is available. In fact, results from such qualitative studies can be important to guide experimental design and instruct the acquisition of quantitative data.

The two case studies presented in this thesis illustrate the multicellular systems biological approach in applications to specific problems in developmental biology, ranging from gene expression to tissue shape and from qualitative mathematical modeling to quantitative computer simulation. Whereas the first study investigates the crosstalk between cell fate regulation and pattern formation in the pancreas, the second explores the interplay between signaling and morphogenesis in the context of vascular network formation. And while the first study is a theoretical study based on qualitative information, in the second case study, an initially qualitative study on is experimentally validated and extended to obtain a quantitative computational model. Although these case studies merely probe the range of possible interplays that are of interest, the in-depth investigation of these disparate topics in developmental biology is intended to highlight the wide variety of approaches and applicability of a multicellular systems biology of development.

### 1.3 OVERVIEW OF THESIS

Figure 1.3 depicts the rationale and structure of this thesis. Developmental processes can roughly be categorized into three developmental mechanisms, cell-face specification, pattern formation and morphogenesis (fig. 1.3A), for each of which well-established modeling methods exist (fig. 1.3B). Interactions between genes can be represented in terms of ordinary or stochastic differential equations to study dynamics of cell fate specification. Interactions between signaling molecules that drive pattern formation can be studied using reaction-diffusion systems. And morphogenetic mechanisms can be captured in a range of cell-based modeling methods that specify the mechanical and behavioral interactions between cells.

Understanding the interactions and feedbacks between these developmental mechanisms, however, necessitates the integration of these modeling methods into a single multiscale modeling framework. In this thesis, such a framework is presented (figure 1.3C). Part I introduces computational methods in multicellular systems biology of development and a new software environment that implements these methods. These methods are applied to two case studies (figure 1.3C) that highlight the diversity of approaches and applications in developmental biology. Part II describes a study on the crosstalk between cell fate regulation and pattern formation in the pancreas and part III describes the interplay between signaling and morphogenesis in the context of vascular network formation. Each part contains a separate introduction describing the biological background and introduces existing modeling approaches (chapters 2, 4 and 7) and are briefly summarized here.

In part I, methods and tools for multi-scale modeling of multicellular systems are introduced that are used in the research work in the following parts. In particular, I present the modeling and simulation environment Morpheus that facilitates the simulation and integration of cell-based models with model describing intra- and extra-

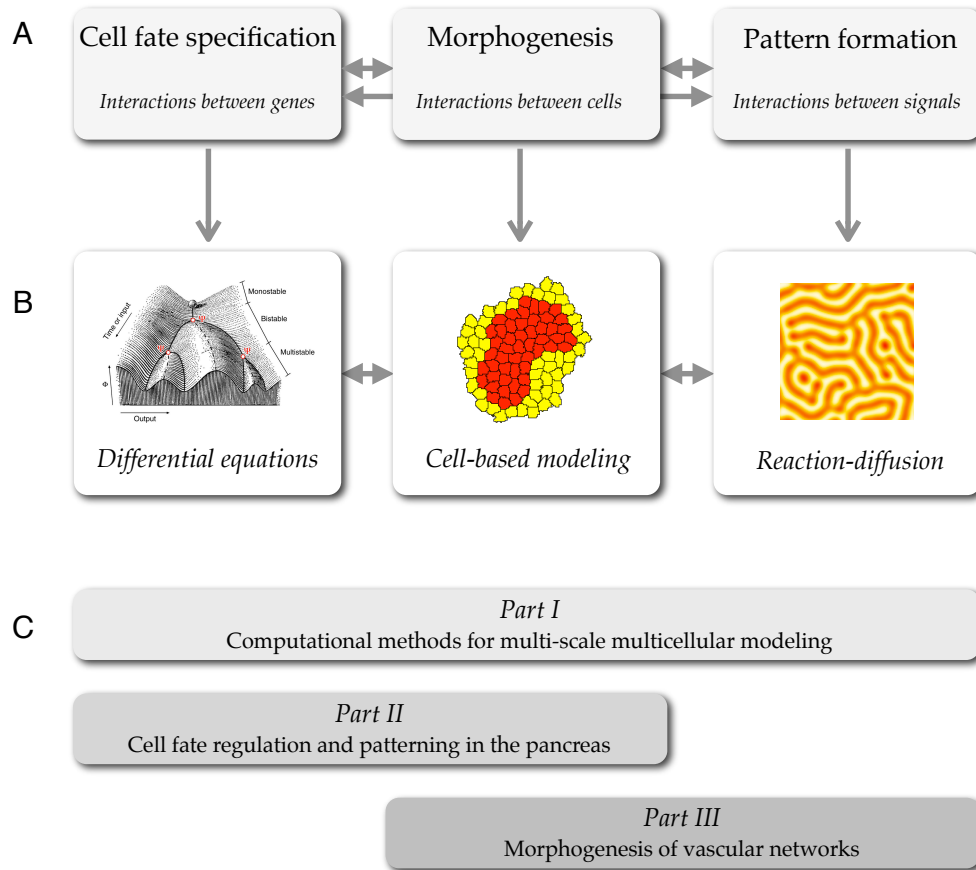


Figure 1.3.: **Rationale and structure of thesis.** (A) Development is driven by interplay between developmental mechanisms and requires the study of interactions at multiple scales. (B) Existing mathematical and computational modeling methods for each of these developmental mechanisms must be integrated to investigate how feedbacks between these developmental mechanisms drives tissue formation. (C) This thesis is divided in three parts. Part I describes the computational methods and software implementation that enables the computational study of interplay between development mechanisms. In Part II, these methods are applied to study cell fate regulation and pattern formation in pancreatic development. Part III describes two applications to the problem of the morphogenesis of vascular networks.



cellular processes. This publicly available software has been designed with a strong focus on usability to facilitate biological experts, even without computational expertise, to use and apply the multicellular systems biological approaches to their own problem domain.

Part II deals with theoretical approaches to understand the regulation of cell fates and spatial patterning. Although both topics have been extensively studied separately, the recent developments in the field of stem cell biology and cellular reprogramming has sparked new interest in mathematical and computational modeling of gene regulatory networks and cell-cell communication and the interplay between these processes. Specifically, I present two theoretical studies on cell fate decision, patterning and reprogramming of pancreatic cells. In the first, a mathematical model of crosstalk between intercellular signaling mechanisms is developed that reproduces the establishment of the scattered spatial distribution and cell type ratio of nascent exocrine and endocrine observed during pancreatic organogenesis (chapter 5). In the second study, this model is extended to provide a theoretical framework to understand experimental observations of *in vitro* acinar-to-islet cell reprogramming and transdifferentiation (chapter 6).

In part III, attention is turned to the interplay between intercellular signalling and morphogenesis during the *de novo* formation of vascular networks. Although previous theoretical work has shown that networks can arise in a variety of ways, including autocrine chemotaxis, some of the assumptions are not substantiated by biological evidence. Two studies are presented that argue that the formation of vascular network formation occurs through paracrine signaling. In the first, it is shown through mathematical and computational modeling that morphometrically realistic vascular networks can arise by paracrine signaling, under the assumption of the retention of the signaling factors by extracellular matrix molecules (chapter 8). In the second study, the predictions generated by this model are experimentally validated *in vitro*. Biophysical measurements obtained from this assay are then used for quantitative modeling to confirm the establishment of network patterns under the paracrine signaling model (chapter 9).

These case studies are not intended to exhaustively explore the range of possible interplays between developmental mechanisms, but to illustrate the applicability of the multicellular systems biological approach from gene expression to tissue shape and from qualitative mathematical modeling to quantitative computer simulation.

#### 1.4 SUPPLEMENTARY ONLINE MATERIAL

Videos and simulation models of discussed computational models discussed in this thesis are available as supplementary online material. Videos are encoded in MP4 format and can be viewed online in a modern HTML5-enabled web browser. Simulation models are encoded in XML-based modeling language MorpheusML (section

3.4). Their simulation requires the Morpheus modeling and simulation environment (chapter 3). Installers for this software are available for MS Windows, Mac OSX and Linux. The supplementary material can be accessed under: <http://walter.deback.net/thesis>.

Part I

COMPUTATIONAL METHODS

---

## INTRODUCTION

---

After establishing all equations and constants for their Nobel Prize winning work on modeling the action potentials in a squid giant axon, Alan Hodgkin and Andrew Huxley used a *Brunsviga 20* mechanical calculator to iterative solve the propagating action potential. Back in 1951, calculating a few millisecond of the differential equations took many weeks and many thousands of rotations<sup>49</sup>. Several years later, the action potential and pacemaker rhythm in cardiac cells was computed on an early electronic computer automatically, but this required providing instructions in endless gibberish of machine code<sup>50</sup>. Now, highly detailed kinetic models of cellular signaling can be computed on histo-anatomically detailed geometric models of the whole heart using massively parallel high-performance computers. These computational models are providing new insights into basic cardiac function and arrhythmia<sup>51</sup> and are already used for drug development<sup>52</sup>.

This remarkable progress in cardiac modeling demonstrates that increases in computational modeling and advances in systems biology have gone hand in hand throughout its history. Not only has the increased computing power facilitated the development of large-scale simulations, it has also boosted the acquisition of reliable quantitative data from high-throughput experiments and microscopy images that are required for quantitative computational modeling.

Although computing power is necessary, it is not sufficient. At least as important are the establishment of computational methods to represent relevant biological processes and their implementation in reliable software tools. The construction and implementation of simulation models poses considerable computational challenges, in particular in cases where the biological system under investigation extends over different levels of biological complexity, from genes, via cells to tissues, each of which are associated with different modeling formalisms. The integration of discrete and continuous, spatial and nonspatial, deterministic and stochastic models is a numerically complicated and error-prone process. While many scientists develop their own custom software for specific research goals, the development of dedicated scientific software based on established software engineering practices becomes increasingly important and offers advantages concerning reproducibility, (re)usability, maintainability and extensibility<sup>53,54</sup>. Moreover, by making the numerical details of computational modeling

transparent, such software also allows less computer-savvy biological experts to use computational modeling to formulate and test hypotheses. It thereby contributes to closing the gap between researchers with expertise in computational modeling and biological experts with extensive domain knowledge.

In this chapter, I introduce the key computational approaches and methods used in multicellular systems biology, I present the state-of-the-art in software tools that implement these methods, and identify some major challenges in this field. The following chapter (3) then focuses on our own computational modeling environment, Morpheus<sup>55</sup>, and describes how this software meets these challenges.

## 2.1 APPROACH

The formation and maintenance of tissues involves genetic regulation, intercellular signaling, external molecular gradients, biomechanical interactions between cells as well as the interplay between these processes. To account for these processes and their interplay, modeling of multicellular systems requires a multitude of mathematical models and computational methods describing biological processes at different spatiotemporal scales. To study the interactions between processes at these various scales, these models are coupled into so-called multiscale models.

Yet, the concept of multiscale modeling has been used in many different ways which confuses the discussion of the challenges involved in the construction and implementation of such models. It is therefore useful to outline its meaning within the context of multicellular systems biology and to describe an approach to avoid the inherent complexities in constructing multiscale models.

### 2.1.1 Multiscale modeling

Biological organisms span an enormous range of spatial and temporal scales, spanning from the molecular length scale ( $10^{-9}$  m) to the length of an organism (1 m) and from the time scale of molecular interactions ( $10^{-3}$  s) to a human lifespan ( $10^9$  s). In many cases, understanding the development or (dys)function of biological organisms requires the integration of data and models across multiple temporal and spatial scales. Yet, in practice, there are considerable differences in how multiscale models are realized.

In molecular systems biology, with its focus on understanding the dynamics of biochemical networks, multiscale modeling typically refers to the coupling of processes at different timescales, since spatial relations are generally not represented explicitly in these models. Coupled dynamics of fast and slow processes, possibly spanning multiple orders of magnitude, can readily be represented in terms of differential equations. However, numerically solving such systems of differential equations is problematic, as they often lead to so-called stiff systems, for which infinitesimally small time

steps would need to be chosen to compute accurately using standard (explicit) solvers. Therefore, to avoid inaccuracy and improve computational efficiency, the model should either be reformulated to reduce stiffness, i.e. by quasi-steady state assumptions, or specific (implicit) solvers must be used that can deal with stiff systems<sup>56</sup>. Apart from these numerical issues, multiscale modeling in molecular systems biology also faces the computational problem of coupling existing models representing dynamics at the intracellular scale to dynamics at the whole-body scale, e.g. linking metabolic network models to physiologically-based pharmacokinetic (PBPK) models. Here, the key challenges are to define relevant points of information exchange between pre-established models, to efficiently compute large numbers of differential equations over long periods of time, and to provide user interfaces aiding the construction of complex models<sup>57</sup>. Yet, multiscale models in this field are typically restricted to ordinary and stochastic differential equation model and do not involve coupling of disparate modeling formalisms.

The concept of multiscale modeling is quite different in the physiology-oriented branch of systems biology. Here, it does not necessarily refer to the coupling of processes at different time scales. Rather, it refers to the coupling of processes at different levels of biological organization<sup>58</sup>: molecules, organelles, cells, tissues, organisms, etc. Note that this explicitly includes processes at different spatial scales. From this perspective, a model is multiscale if and only if it includes submodels representing two or more processes at different levels of organization. Because the processes at each level typically represented using specific modeling formalisms, it follows that multiscale modeling involves the integration of different modeling methods. Therefore, the issue of multiscale modeling becomes a problem of providing a reliable computational infrastructure to couple radically different modeling methods, e.g. discrete and continuous, spatial and nonspatial, deterministic and stochastic models. The large diversity of processes that need to be coupled to understand organ physiological and the number of available modeling methods to represent them aggravates this problem. As a consequence, efforts to enable such multiscale modeling in this field are focused on the establishment of new standard formats to flexibly represent spatial<sup>59</sup> and mechanical<sup>60</sup> models as well as the development of software tools to simulate these models.

#### *Middle-out modeling*

Multiscale models can be constructed in various ways. One can start at the molecular level and progressively include representations of spatial interactions at the cellular, tissue level. Or one can start by describing functionality at tissue level and work down towards its implementation at the cellular and molecular levels. However, both approaches, bottom-up and top-down, have their particular problems. Whereas bottom-up approaches easily become computationally intractable due to the large number of components even at the molecular level, top-down approaches often remain phenomenological descriptions rather than mechanistic models. A pragmatic alternative

to multiscale modeling is offered by the middle-out approach<sup>39,61,62</sup>. In this approach, one particular level of organization is adopted as the main level of abstraction, depending on the relevance to the phenomenon as well as the available data. Subsequently, processes at lower and higher levels are explicitly included and coupled to the existing model, if deemed relevant to the problem under investigation. This approach avoids the pitfalls of (i) the sheer complexity of describing biological function from the bottom-up and (ii) the mere phenomenological description of function in the top-down approach. Moreover, the middle-out approach does not presuppose a certain direction of causality in biological regulation and is therefore unbiased with respect to the study of regulation through feedback between levels of biological organization.

Key questions in multicellular systems biology concern the interaction between different levels of biological organization. Therefore, models in this field often consist of various submodels for processes at different biological levels. Hence, they fall under the second interpretation of the term 'multiscale'. As a method to retain tractability in the construction of such models, a middle-out modeling approach is adopted. Here, the cellular level is taken as the main level of abstraction and the central point of integration, whereas models of intracellular regulation and tissue-level extracellular processes are included as deemed necessary.

## 2.2 COMPUTATIONAL METHODS

When applied to common problems in multicellular systems biology, the multiscale, middle-out approach described above results in a conveniently small number of modeling methods that can be re-used in various combinations. The most suitable combination depends on the biological context as well as the research question. Figure 2.1 shows the most commonly used combination of formalisms. Intracellular dynamics are described by ordinary differential equations (fig. 2.1C), the spatial distribution of extracellular signaling molecules are described by reaction-diffusion equations (fig. 2.1A) and cellular behavior and intercellular mechanics are modeled in one of several cell-based model formalisms (fig. 2.1B).

### 2.2.1 Dynamic modeling

Describing the dynamics of interaction between biochemical components is at the heart of systems biology. The most frequently used tool to model these dynamics are coupled sets of ordinary differential equations (ODEs). Hodgkin and Huxley, for instance, used ODEs to describe the membrane current as an interaction between the potassium and sodium ion channel activation<sup>63</sup>. ODEs are now ubiquitously used in systems

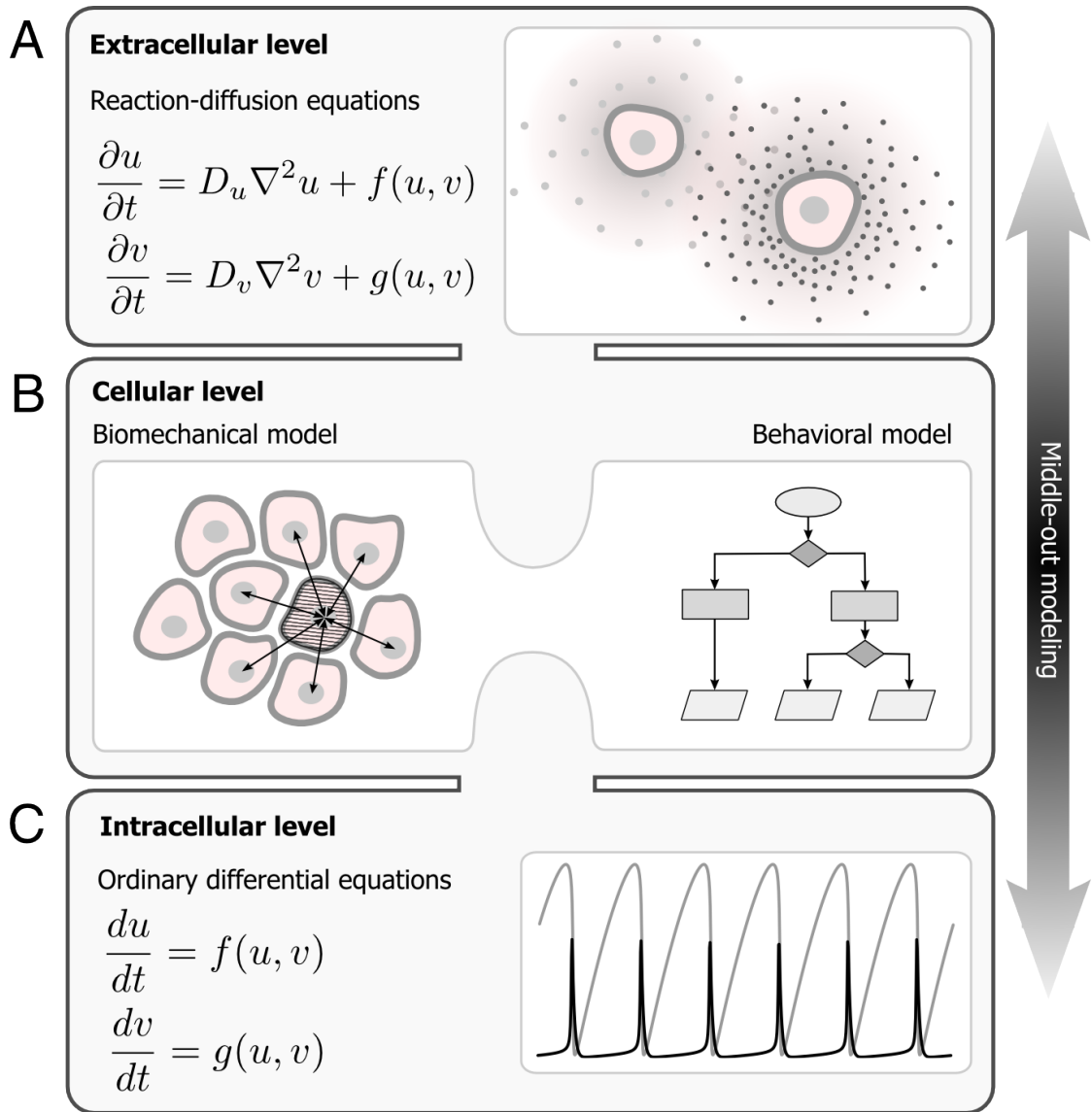


Figure 2.1.: **Computational methods in multicellular systems biology.** Models consist of the following components that can be coupled using a middle-out strategy starting at the cellular level. (A) Reaction-diffusion equations are used to represent spatial distributions and interactions of signaling agents. (B) A variety of cell-based models can be used to represent cellular biomechanics (see figure 2.2) and rules specifying cell behavioral decisions. (C) These cellular aspects can be controlled by ordinary or stochastic differential equations that represent intracellular regulatory dynamics. (Figure adapted from W. de Back, T. Sütterlin, A. Deutsch and N. Grabe, in preparation.)



biology to describe the dynamics of biochemical networks, including metabolic networks<sup>64</sup>, signaling pathways<sup>65</sup> and genetic regulatory networks<sup>66</sup>.

ODEs represent the behavior of biological systems in terms of continuous-time non-linear dynamical systems. A large number of mathematical and computational tools are available to simulate and analyze such systems, ranging from time course simulation, stability and bifurcation analysis to parameter identifiability and optimization techniques. Moreover, there are a number of well-established software tools available for their simulation and analysis, such as COPASI<sup>67</sup> and CellDesigner<sup>68</sup>.

Ordinary differential equations assume that the modeled system is composed of homogeneous, well-mixed populations of components that behave deterministically and instantaneously. The assumption of homogeneity can be relaxed using compartmentalized models, while stochastic and delay differential equations (SDEs and DDEs) can be used to relax the latter two assumptions, although at the expense of mathematical tractability.

Dynamic models of intracellular regulation can be linked to cell-based models to serve to control the cell's mechanical properties or behavioral decisions (fig. 2.1). For instance, intercellular adhesion may depend on dynamic models of cadherin expression, or the propensity for cell division may depend on an intracellular model describing cell cycle progression.

### 2.2.2 Spatial modeling

While the behavior of molecular regulatory networks can often be faithfully studied in non-spatial models, studying the causes and effects of spatial heterogeneity requires models that explicitly account for spatial interactions. The emergence of patterns and shapes is, of course, of particular interest to developmental biology. Gradient models like Wolpert's French flag model, for instance, presented a conceptual mechanism allowing patterning of cells in a concentration-dependent manner when exposed to a spatial gradient<sup>69</sup>. Even earlier, Alan Turing has shown that such gradients can arise by self-organization, given the right interactions between diffusive morphogens formation<sup>44,70</sup>. Since then, partial differential equations, and in particular reaction-diffusion equations, have been used to study many examples of pattern formation in biological systems<sup>24</sup> ranging from animal coat patterns<sup>71</sup> to vascular network formation<sup>72</sup> and from somitogenesis<sup>73,74</sup> to cell polarization<sup>75,76</sup>.

In the context of multicellular systems biology, reaction-diffusion equations and gradient models are coupled to cell-based models to describe the diffusion of signaling agents through the tissue as well as (un)-binding to extracellular matrix components (fig. 2.1). Although reaction-diffusion equations are also used to describe the spatiotemporal behavior of populations of cells at a macroscopic level, this is typically not appropriate to represent the heterogeneity of cellular behaviors at the level of the single cell.

### 2.2.3 Cell-based modeling

To model growth, motility and patterning in heterogeneous cellular populations, cells are represented as discrete entities or agents. A range of computational methods have been developed to describe the interactions between cells as discrete entities. These cell-based models describe cells with respect to cell shape, motility, growth and mechanical interactions in various levels of detail<sup>77</sup>. As shown in figure 2.2, they can be categorized in on- and off-lattice models, depending on whether cells are spatially represented in continuous space or discretized on a lattice. Alternatively, they can be classified according to the way in which systems dynamics is calculated, either based on rules, forces or free energy.

In lattice-based methods such as cellular automata<sup>78</sup> and the cellular Potts model<sup>79</sup>, cells are spatially represented as occupying one or more lattice sites. In classical cellular automata, cells are immotile and the state of cells depends on local interactions with adjacent cells, although extensions such as lattice gas cellular automata also allows the description of state-dependent cell motility<sup>78</sup>. The cellular Potts model describes cells as connected domains of lattice sites where cell motility arises by local spin-copy events minimizing a free energy function that describes mechanical cell properties such as limited compressibility, contractility and cell-cell adhesion<sup>79</sup>.

Off-lattice models such as cell-center models<sup>80,83</sup>, vertex models<sup>81,84</sup> and subcellular elements model<sup>82</sup> describe cells as a single or collection of points in continuous space. In cell-center models, point-like cells are connected by springs to model adhesion forces that are balanced by a viscous drag term. Vertex models represent cells as polygons whose vertices are shared by multiple cells. Cell rearrangements are the result of movements of these vertices according to equations of motions that may be based on a free energy function<sup>81</sup>. The subcellular element models provides a more fine-grained model of cell shape by representing each cell as a network of connected elements. Cell shape and motility evolves as a result of a balance between intra- and intercellular potentials<sup>82</sup>.

Each of these computational modeling methods have their strengths and weaknesses. To focus on the weaknesses, the cellular Potts model is essentially an equilibrium rather than a dynamic model<sup>85</sup>, the centre-based model does not account for cell shape, the vertex model is limited to represent confluent epithelial tissues in 2D and, due to high computational costs, the subcellular element model is impractical to represent large tissues. Therefore, none of these cell-based formalisms has been widely adopted as a standard model representation for biological cells. And, unfortunately, studies with critical analyses of cell-based model behavior and systematic comparisons between model formalisms are still rare<sup>85,86</sup>. The choice for a particular cell-based model formalisms thus typically depends on the specific application domain and the availability of data on cell shape and biophysical parameters. Despite the lack of a standard

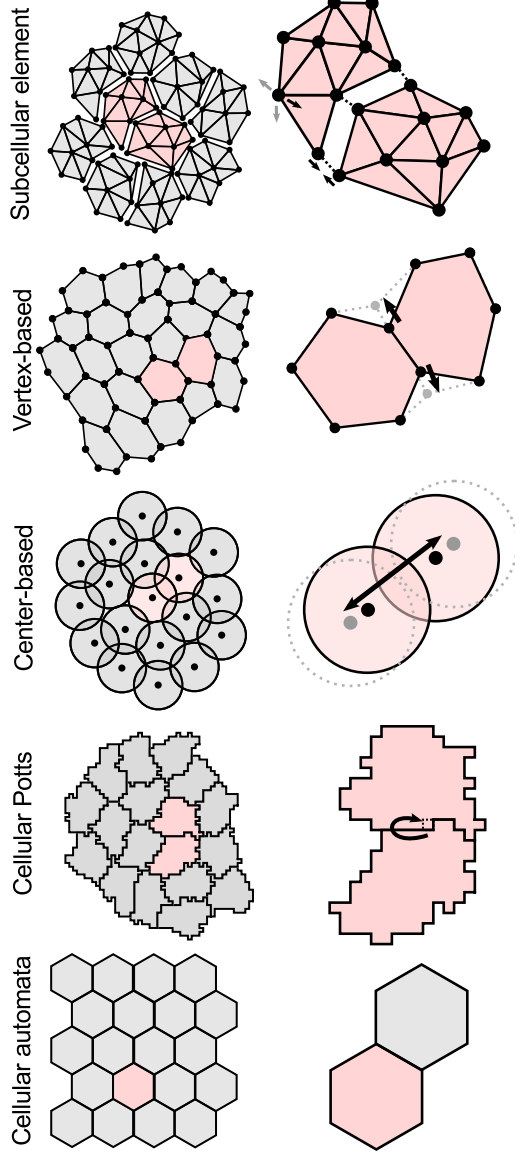


Figure 2.2.: **Cell-based model paradigms.** Computational methods to describe cells in varying level of spatial and mechanical detail: Cellular automata (CA)<sup>78</sup>, cellular Potts model (CPM)<sup>79</sup>, center-based models (CBM)<sup>80</sup>, vertex models (VM)<sup>81</sup> and subcellular elements models (SEM)<sup>82</sup>. The two leftmost methods (CA and CPM) represent cells as sites or domains on a regular lattice, while the others (CBM, VM and SEM) represent cells in continuous space. Whereas the CA and CBM describe cells as point particles, the remaining methods explicitly include cell area and shape in various level of detail. Cell motility is either described indirectly in terms of a free energy to be minimized (CPM, CBM, VM) or directly through explicit equations of motion (CBM, VM, SEM). (Figure adapted from W. de Back, T. Sütterlin, A. Deutsch and N. Grabe, in preparation.)

representation of cells, cell-based modeling methods in general have been established as a valuable tool for the investigation of the behavior of multicellular systems.

While all the aforementioned computational modeling methods are well-established by themselves, their combination opens up new possibilities to understand the interplay between developmental mechanisms. Yet, their integration provides several computational challenges. It demands the implementation of numerical solvers for all or a subset of the above modeling methods as well as methods to integrate these into reliable multiscale models. Consider a simple model in which the propensity for cell division in a cell-based model depends on an intracellular model describing the cell cycle. Even this simple example requires the coupling between discrete and continuous models, deterministic and stochastic models, and spatial and non-spatial models. Mathematical and numerical methods to tightly integrate such models are not yet well developed and the requirements differ between specific modeling formalism, rendering the construction and implementation of multiscale multicellular models a challenging task.

### 2.3 SOFTWARE ENVIRONMENTS

The computational requirements for multiscale multicellular modeling go well beyond the capabilities of standard software for systems biology. Therefore, researchers are forced to implement their own computational models by hand using a general-purpose programming language or generic software such as MatLab. However, manual implementation has several drawbacks since (1) it requires considerable expertise on mathematics, biophysics and software design, (2) it transforms the task of biological modeling into a programming problem that precludes most biologists and (3) it hampers the reproducibility of results by third parties.

Recently, a new generation of software environments has emerged that facilitate the simulation of multiscale models of multicellular systems (see table 2.1). These software tools all have similar functionality in the sense that they allow users to construct and simulate multiscale models by linking cell-based models to models of intra- and extracellular dynamics. Despite the fact that all environments implement the common components depicted in figure 2.1, there are significant differences between them.

One of the key differences is the support for cell-based modeling paradigms (see table 2.1). Most platforms are specialized on a particular cell-based model and are therefore subject to the specific strengths and weaknesses of that method. Only few support the simulation of multiple cell-based formalisms. Apart from improving modeling flexibility, this also enables the study of model behavior under various assumptions on biomechanical and cell shape properties. Chaste, for instance, provides implementations for center-based, cellular Potts and vertex models, and is therefore well-suited for comparative studies using these different methods<sup>86</sup>.

Software	Cell-based model(s)					Modeling language
	CA	CPM	CBM	VM	SEM	
Biocellion <sup>87</sup>			•			C++
CellSys <sup>88</sup>			•			C++
CompuCell3D <sup>89</sup>		•				Python, XML, C++
Chaste <sup>90</sup>		•	•	•		C++
EPISIM <sup>91</sup>	•		•			Process diagrams
GNOMO					•	Fortran
Morpheus <sup>55</sup>	•	•				MorpheusML
VirtualLeaf <sup>92</sup>				•		C++

Table 2.1.: **Software environments for multicellular systems biology.** New generation of modeling and simulation software that implements computational methods show in figure 2.1. Cell-based models: CA=cellular automata, CPM=cellular Potts model, CBM=center-based model, VM=vertex model, SEM=Subcellular element model.

Another major difference between the platforms lies in the languages and interfaces that they provide to the user to construct their models. All software platforms hide the numerical details of the implementation by exposing their functionality in terms of a high-level language. Yet, these vary from application programming interfaces (API) in general purpose languages such as C++ and scripting language such as Python to domain-specific declarative languages or even graphical modeling interfaces. Whereas modeling in general purpose programming languages offers maximum modeling flexibility and computational performance, the use of biological and mathematical terminology in domain-specific languages and graphical modeling tools typically provide improved usability and may be more accessible to biological users without extensive computational knowledge.

### 2.3.1 Challenges

The key functionality of these software environments is the flexible integration of modeling formalisms. However, a major challenge for all environments is the optimization of usability without sacrificing modeling flexibility. While flexibility and extensibility are essential to facilitate innovative modeling studies, usability is crucial to widen the target audience from computational experts to biological researchers with extensive domain knowledge. Moreover, usability is important to allow integration of modeling and simulation into the everyday workflow in biology, similar to e.g. image analysis and statistical software.

Another formidable challenge is the establishment of standardized formats for the representation, storage and exchange of models of multicellular systems. Standard modeling formats, such as the systems biology markup language (SBML)<sup>93</sup>, have gained a prominent place in systems biology where they facilitate the open exchange between users and software. Yet, no such standards exist that are able to represent multiscale models of multicellular systems. On the one hand, this is complicated by the fact that the establishment of standard requires a concerted community effort. On the other hand, it is difficult to represent dynamic multiscale models encompassing multiple modeling formalisms in a static declarative markup language<sup>94</sup>.

---

## MORPHEUS: A MODELING ENVIRONMENT FOR MULTISCALE AND MULTICELLULAR SYSTEMS BIOLOGY\*

---

### 3.1 INTRODUCTION

Computational modeling crucially depends on software. In contrast to the modeling of molecular reaction networks, for which a large and diverse collection of software tools is available<sup>†</sup>, software that supports computational modeling of multiscale multicellular systems is scarce. Therefore, many researchers are forced to write their own custom-made software, which often leads to computational models that are irreproducible, not extensible, poorly documented and difficult to exchange between researchers.

Recently, a number of dedicated software platforms have come available<sup>55,88–92</sup> that provide reliable and reusable implementations of established cell-based models that can be coupled to models of intracellular and extracellular dynamics, rendering them suitable to investigate tissue dynamics over multiple spatiotemporal scales. However, most of these platforms are targeted at computational experts and require in-depth knowledge of programming, while the few software platforms that provide intuitive graphical interfaces, lack the flexibility to be customized for innovative modeling studies.

We have designed a modeling and simulation environment for multiscale and multicellular systems, Morpheus, to provide reusable implementations of computational methods that can be flexibly combined into complex multiscale simulations. Models are constructed from within a user-friendly graphical user interface, that does not require programming expertise. Morpheus facilitates the simulation of a range of computational modeling formalisms including ordinary, stochastic and delay differential equations, Boolean models, rule-based models, cellular automata, cellular Potts models, gradient-based models and reaction-diffusion models. These formalisms can be

---

\* This chapter includes text and figures from the publication and supplementary material of: Jörn Starruss, Walter de Back, Lutz Brusch and Andreas Deutsch, *Bioinformatics*, 30(9):1331–1332, 2014. Author contributions: Jörn Starruss and Walter de Back conceived, designed and implemented the software. Jörn Starruss is core developer and Walter de Back is contributing developer. Walter de Back wrote the paper.

† The SBML software matrix ([http://sbml.org/SBML\\_Software\\_Guide](http://sbml.org/SBML_Software_Guide)) mentions 274 software packages that support the simulation and analysis of biochemical network models written in SBML.

used in isolation or combined to represent multiscale and morphodynamic systems. It is available as a ready-to-use application with graphical user interfaces (GUI) for modeling and simulation<sup>55</sup> and it provides a well-documented framework with a C++-based plug-in architecture that allows customization and extensibility.

A key technology that enables Morpheus to combine modeling flexibility with usability is a novel XML-based model description language for multiscale multicellular models, MorpheusML. This language enables a separation of concern between modeling and implementation in which the model, created in the GUI, fully specifies what should be simulated and the simulator determines how this is accomplished algorithmically. Similar to markup languages such as the systems biology markup language (SBML)<sup>93</sup>, this language combines biological terminology with symbolic mathematical expressions to represent model dynamics. However, it adds the possibility to represent the dynamics of spatial and cell-based models.

Internally, models in MorpheusML format are interpreted by the simulator and executed as a series of plugins. Plugins can be added or customized to meet specific modeling goals. The order and update frequencies of the various plugins are automatically derived from the dependencies and spatiotemporal contexts of the symbols specified in the model such that the correctness and efficiency of the simulation are ensured. This scheduling automates the error-prone process of model integration between different model formalisms.

Together, the user-friendly user interface, the model description language and the extensible simulation framework provide a powerful toolset that allows computational modeling of multicellular system to be integrated in the common workflow in biology, comparable to software for e.g. image processing or statistical analysis.

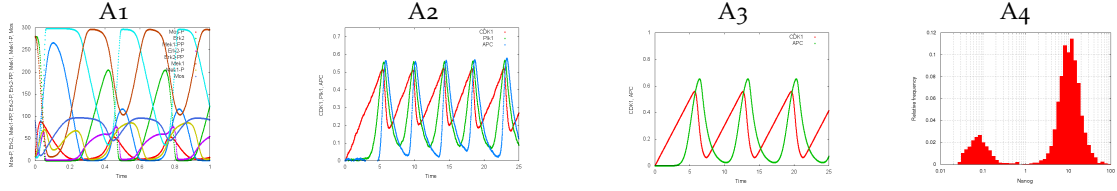
This chapter provide an overview of the versatility and usability of Morpheus. We first describe the modeling formalisms that are available and show the versatility of type of models that can be configured. We then show how to use construct, configure and execute models using the graphical user interface. To understand how models are stored, interpreted and executed, we describe the model description language MorpheusML. Finally, it is shown how Morpheus performs automatic model integration and how its plugin architecture supports extensibility.

### 3.2 MODELING FORMALISMS

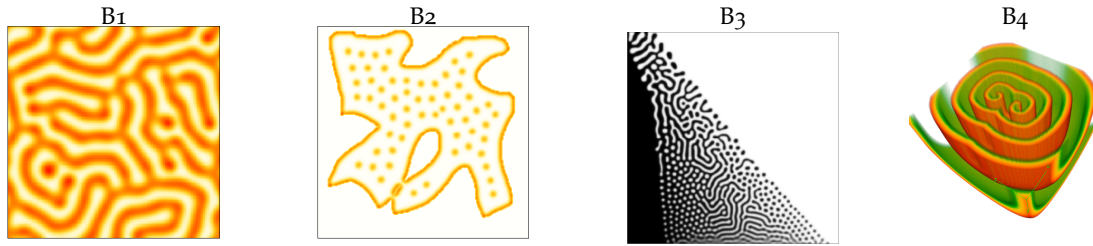
Morpheus supports the simulation of a number of modeling formalisms based on differential equations as well as several cell-based models. These core modeling formalisms can be used in isolation or combined into multiscale models. Moreover, the use of mathematical expressions such as functions, rules and discrete events, allows for a versatile array of modeling methods, as illustrated in table 3.1.



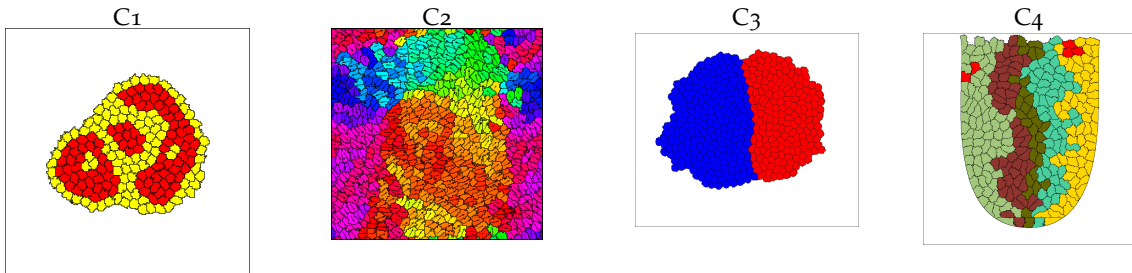
## Differential equations models



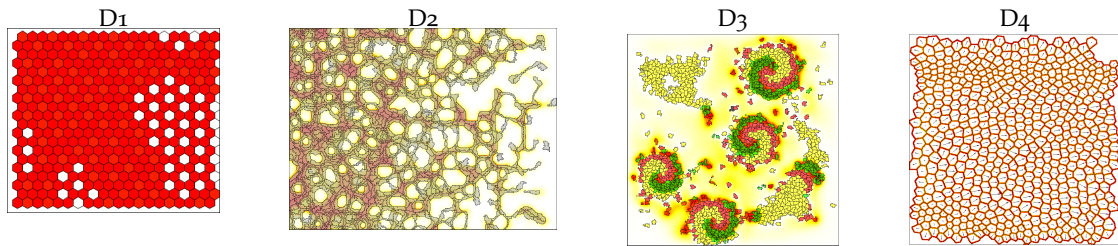
## Reaction-diffusion models



## Cellular Potts models



## Multi-scale models



## Miscellaneous

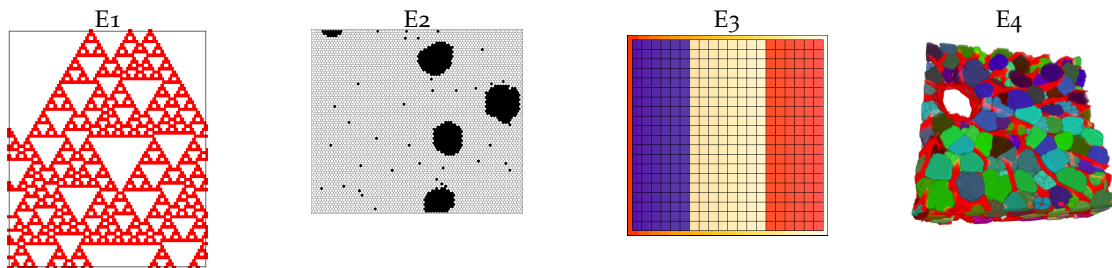


Table 3.1.: Use cases of different modeling formalisms in Morpheus. (Caption on next page)

Figure 3.1.: (A1) Ordinary differential equation (ODE) model of MAPK cascade<sup>95</sup>, imported from SBML format via BioModels. (A2) Stochastic differential equation (SDE) model of *Xenopus* cell cycle<sup>65</sup>. (A3) Delay differential equation model (DDE) of cell cycle<sup>65</sup>. (A4) ODE population model of Nanog expression in mouse embryonic stem cells<sup>96</sup>. (B1) Gierer-Meinhardt model with short-range activation and long-range inhibition<sup>70</sup>. (B2) Gierer-Meinhardt model in arbitrary domain, imported from image<sup>70</sup>. (B3) Spatially varying parameters in Turing model<sup>97</sup>. (B4) 3D simulation of Barkley model of excitable media<sup>98</sup>. (C1) Cellular Potts model (CPM) of cell sorting based on differential adhesion<sup>79</sup>. (C2) CPM model of persistent motion in monolayers, adapted from<sup>99</sup>. (C3) CPM adaptation of vertex model of boundary formation, adapted from<sup>100</sup>. (C4) CPM model of clonal growth of stem cells in intestinal crypt. (D1) Lattice of spatially coupled SDE representing cell fates in pancreas<sup>101</sup>. (D2) Multiscale CPM / reaction-diffusion model of vascular network formation with varying cell density<sup>102</sup>. (D3) Multiscale model of *Dictyostelium* aggregation with cAMP wave propagation leading to aggregation of chemotactic cells, adapted from<sup>103</sup>. (D4) Planar cell polarity model with reaction-diffusion model<sup>76</sup> on membranes, image of segmented cells imported from<sup>104</sup>. (E1) 1D cellular automata (CA) model (rule 30) of patterning of sea shells<sup>105</sup>. (E2) Stochastic interacting particle system (IPS) of aggregation. (E3) Wolpert's French flag model of positional information<sup>69</sup>. (E4) 3D cell shapes imported from TIFF images (image courtesy of Zerial lab, MPI-CBG). Videos and simulation models in MorpheusML format are available in the Supplementary Online Material under <http://walter.deback.net/thesis>.

### 3.2.1 Differential equations

Morpheus supports the simulation of ordinary, stochastic and delay differential equations as well as reaction-diffusion systems, see figure 3.2. Differential equations are perhaps the most widely used formalism in systems biology and have been applied to all kinds of biological networks, from signaling pathways and metabolic networks to gene regulatory networks<sup>106</sup>. Systems of coupled sets of non-linear differential equations can be used to describe the rate of change of continuous variables. In systems biology, this is typically used to describe the concentration of molecular species as a function of time.

One can perform time-course simulations and predict responses to different stimuli, if a fully detailed kinetic model can be constructed. Accurate simulation of system dynamics, however, requires detailed knowledge of the involved reaction mechanisms as well as kinetics parameters for which experimental measurements may be difficult to obtain. But even incomplete knowledge of reaction kinetics, differential equations form an attractive modeling formalism due to the fact that, at least small systems of ordinary differential equations (ODEs) can be analysed using well-established methods such as phase diagrams (revealing qualitative changes in state space), linear stability analysis (showing stability properties of equilibria) and bifurcation diagrams (relating stability properties to changes in parameters).

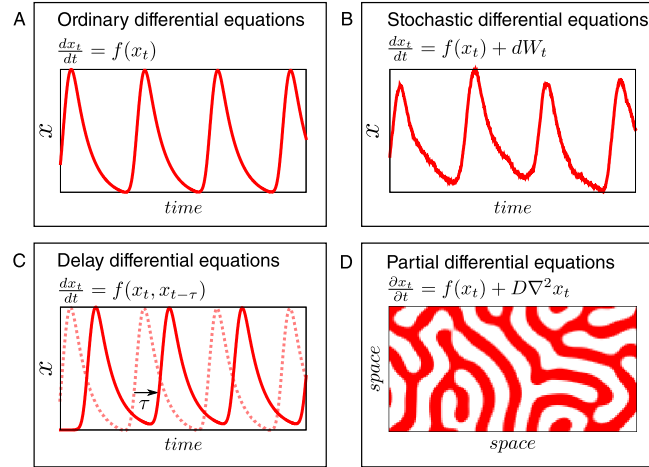


Figure 3.2.: **Different types of differential equations available in Morpheus.** (A) Deterministic ordinary differential equations are entered in conventional in-fix notation. (B) Stochastic differential equations can be simulated by including a normally distributed stochastic term, `rand_norm([mean], [stdev])`. (C) Delay differential equations are simulated by assigning to the special `DelayProperty` with a given delay  $\tau$ . (D) Partial differential equations of the reaction-diffusion type can be simulated by assigning to a `Layer`, defining a value field including a diffusion coefficient.

Stochastic differential equations (SDEs), figure 3.2B, are differential equations that include a stochastic term to represent randomness inherent in the thermal nature of chemical reactions. Whereas deterministic ordinary differential equations, figure 3.2A, are based on the assumption that random fluctuations are averaged out in a large population, in small population, such as transcription factors binding to one or two promoter regions, stochastic effects may persist. Stochastic effects can play a major role by allowing the system to escape unstable equilibria by random perturbation.

Delay differential equations (DDEs), figure 3.2C, are useful to represent lags in reactions due to e.g. transport delays or to lump together complicated processes and only account for the time required for these processes to occur. Delays can qualitative alter dynamics. For instance, adding delays to systems with negative feedback, can result in oscillatory behavior that is otherwise not observed<sup>65</sup>.

Whereas the differential equation systems above assume spatial homogeneity, partial differential equations (PDEs), in particular reaction-diffusion (RD) systems, can account for heterogeneous spatial distributions<sup>24</sup>. This can be used to model the establishment of morphogen gradients providing positional information or cues for cell motility<sup>102,107</sup> as well as the self-organization of spatial patterns by interaction between diffusible species<sup>44,108</sup>.

*Numerical solvers*

Morpheus implements solvers for initial value (Cauchy) problems for ODEs based on the finite difference methods for different orders (Euler, Heun, Runge-Kutta) and fixed time stepping. Finite difference methods are explicit numerical integration schemes to solve differential equations given initial conditions and allow straight-forward implementation and circumvent the need to algebraically determine a Jacobian, as with implicit methods. However, explicit numerical methods require user-specified time steps to achieve sufficient accuracy of solution and guarantee numerical stability and cannot be used to solve stiff systems that generally occur when dynamics of variables vary over order of magnitude, requiring unacceptably small time steps to ensure numeric stability. Although adaptive time-stepping methods may solve the first problem, Morpheus uses fixed user-specified time steps to enable the simulation schedule and time intervals to be determined at initialization in order to offer automated model integration, as explained below.

For SDEs, the Maruyama method is used automatically whenever stochastic terms are used in differential equations. This method ensures that the noise amplitude is scaled with the user-specified numerical integration time step. Morpheus also supports the simulation of DDEs using a special variable that returns values at time  $t$  that were assigned to it at time  $t - \tau$ , where  $\tau$  denotes a given delay. Note, however, that the delay  $\tau$  is fixed during simulation. Reaction-diffusion systems are solved using the sequential operator splitting method in which the original problem is split into two subproblems (the reaction and diffusion steps) that are solved sequentially, both for the same time step. It uses the central difference method to solve the diffusion equation, based on the diffusion coefficient for each species and the spatial discretization of the lattice. During initialization, the numerical time step for the reaction step is adopted by the diffusion problem and automatically adjusted in order to satisfy the Courant–Friedrichs–Lewy (CFL) condition ensuring numerical stability.

## 3.2.2 Cell-based models

Morpheus also supports the simulation of a number of discrete cell-based models in which cells are spatially represented on a lattice. These include the cellular Potts model and several derived formalisms such as coupled ODE lattices, cellular automata and interacting particle systems.

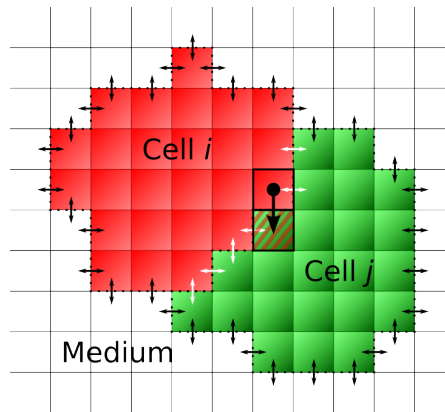


Figure 3.3.: **Cellular Potts model.** Cells are represented as connected domains on a lattice. The area of a cell  $A_\sigma$  equals the number of lattice sites it occupies (colored boxes) and the cell perimeter  $P_\sigma$  is the number of interfaces with other cells or medium (dotted lines). For each interface, a characteristic binding energy  $J$  is defined (small arrows) that may differ for cell-cell (white) and cell-medium (black) interfaces. Cell shape changes and motility arise through spin-copy attempts (big arrow) along the cell membrane mimicking protrusions and retractions. Acceptance of these attempts depends on the associated change in the free energy  $\Delta H$ .

#### *Cellular Potts model\**

The cellular Potts model (CPM), also known as the Glazier--Graner--Hogeweg model<sup>109</sup>, is a modeling framework in which cells are represented as discrete entities that are spatially extended as domains on a lattice. It describes cell and tissue behavior in terms of cell surface mechanics<sup>110</sup>, including volume conservation, adhesion and cortical tension. It is applicable when the details of cellular behavior and intercellular interactions are essentially determined by the shape and the size of the individual cells as well as the length of the contact area between neighboring cells. This model class has originally been developed by Glazier and Graner<sup>111</sup> to study cell sorting, i.e. the observed segregation of heterotypic cell aggregates into spatially confined homotypic cell clusters. The CPM was introduced to explore the tissue-scale consequences of the differential adhesion hypothesis that holds that cell-type-dependent disparities in the expression of molecules that regulate intercellular adhesion are responsible for cell sorting. Since then, this formalism has been elaborated and applied to study a wide range of morphogenetic phenomena in developmental biology<sup>112</sup>.

\* This subsection contains texts adapted from: Anja Voss-Böhme, Jörn Starruss and Walter de Back, Cellular Potts model, In: *Encyclopedia of Systems Biology*, W. Dubitzky and O. Wolkenhauer and H. Yokota and K.-H. Cho (Eds.) Springer, 2013.

### Configuration

In the CPM, a population of  $N$  cells are represented by a connected subdomains on a regular lattice. Although square and hexagonal 2D lattices and cubic 3D lattices can be used, for clarity the description is restricted to a 2D regular square lattice, as depicted in figure 3.3. All lattice sites that a particular cell occupies are labeled by the same index  $\sigma = \{1, 2, \dots, N\}$ , while a special index  $\sigma = 0$  labels the medium, i.e. all lattice sites not occupied by cells. In this formalism a cell has finite volume and deformable shape. The interfaces between two different lattice sites  $x$  and  $x'$  with different indexes  $\sigma_x \neq \sigma_{x'}$  represent membrane boundaries between cells or between cells and the medium. To each of these boundaries, a characteristic binding energy is assigned:  $J_{cc}$  when the interface is between two different cells and  $J_{cm}$  when it lies between a cell and the surrounding medium. An energy penalty increasing with the cell's deviation from a selected target area  $A_\sigma$  imposes an area constraint on the cells. Optionally, an additional constraint on the perimeter of the cell is included by a penalty on deviation from a particular target perimeter  $P_\sigma$ .

The corresponding Hamiltonian is defined as follows:

$$H = \sum_{\{x, x'\}_n} J_{\tau(\sigma_x)\tau(\sigma_{x'})} (1 - \delta_{\sigma_x, \sigma_{x'}}) + \lambda_A \sum_{\sigma > 0} (a_\sigma - A_\sigma)^2 + \lambda_P \sum_{\sigma > 0} (p_\sigma - P_\sigma)^2 \quad (3.1)$$

where  $\tau(\sigma_x)$  represents the type  $\tau$  of a cell  $\sigma$  occupying a grid space  $x$ , which in this case can only be either cell (c) or medium (m). The term  $(1 - \delta_{\sigma_x, \sigma_{x'}})$  with Kronecker delta  $\delta$  ensures that binding energies are only considered between non-identical cells. The terms  $(a_\sigma - A_\sigma)$  and  $(p_\sigma - P_\sigma)$  represent the deviations of the current cell shape from target values and  $\lambda_A$  and  $\lambda_P$  represent the cell's resistance to such deformations. While the first summation is taken for interfaces between the  $n^{\text{th}}$ -order neighbors in each lattice site, the other two sum all cells with the exception of the medium ( $\sigma > 0$ ).

### Dynamics

The CPM is a time-discrete Markov chain where the transition probabilities depend on the Hamiltonian  $H$ . Dynamics are generated by a modified Metropolis algorithm. This algorithm randomly chooses a lattice site  $x_{\text{target}}$  and computes what the difference in energy,  $\Delta H$ , would be if a randomly selected neighboring site  $x_{\text{source}}$  would copy its state into the target site, see figure 3.3. The probability of accepting the change,  $P(\Delta H)$ , depends on the difference in the energy costs:

$$P(\Delta H) = \begin{cases} 1 & \text{if } \Delta H \geq 0 \\ e^{\frac{-(\Delta H)}{T}} & \text{otherwise,} \end{cases} \quad (3.2)$$

such that extensions or retractions of a cell that diminish the free energy  $H$  are always accepted and those that increase  $H$  are accepted according to a Boltzmann distribution. In this way, the shape of cells are locally updated. Parameter  $T$  is a biological analogue to the energy of thermal fluctuations in statistical physics and it is considered here as a measure of cell motility. The unit of time in CPM is usually defined by the number of random update attempts equal to the number of lattice sites, defining one Monte Carlo step.

#### *Extensions and multiscale coupling*

An important advantage of the cellular Potts model is that it can be readily extended in various ways. For instance, one can include terms in the Hamiltonian to represent, for instance, constraints on cell length<sup>113</sup>. Cell division can be modeled by assigning a new cell index  $\sigma$  to half of the lattice sites occupied by a cell, and cell death or apoptosis can be represented by changing a cell's index to medium  $\sigma = 0$  or setting the target area to zero,  $A_\sigma = 0$ .

The CPM can also be coupled to other model formalisms. All cell-based parameters, such as target area  $A_\sigma$  or cell adhesion affinities  $J$  but also e.g. probability for cell division, may depend on submodels that represent intracellular dynamics. In this way, ordinary differential equations modeling signaling pathways or gene regulation can control cellular behavior and tissue morphogenesis in the CPM.

Additionally, cellular behavior may be linked to models of morphogens in the extracellular environment, represented by static gradients or partial differential equations. Non-Hamiltonian terms can be included that alter the probability of accepting updates based, for instance, on the local concentrations of a morphogen gradient to represent chemotaxis. A simple method to model chemotaxis, as used in chapters 8 and 9, is to bias cellular extensions in the direction of higher concentrations by

$$\Delta H_{\text{chemotaxis}} = \Delta H - \mu(c_{\text{target}} - c_{\text{source}}) \quad (3.3)$$

where the free energy is altered proportionally to the local difference in morphogen concentration  $c$  at the target and source sites and the chemotactic response parameter  $\mu$ <sup>103</sup>.

#### *Critics*

Some have criticized the CPM framework with respect to the fact that some parameters, such as the temperature  $T$ , are difficult to relate to measurable biophysical quantities. However, work by Hogeweg and colleagues have shown how CPM parameter can be mapped to biological and physical properties of cells<sup>114</sup> and, more recently, how the CPM parameters can be re-scaled to other cell surface mechanics models such as the vertex model (Stan Maree, personal communication). In work included in this

thesis, CPM parameters involving chemotactic sensitivity have been estimated using experimental measurements of cellular motility in a microfluidic device (see chapter 9)<sup>115</sup>.

Others have pointed out that the CPM, due to its method of energy minimization, is an equilibrium model and that observed dynamics during the process of energy minimization do not reflect true dynamics in the mathematical sense but merely relies on apparent similarity to biological phenomena<sup>85</sup>. Although this is true for CPM per se (as well as for other formalisms based on Metropolis kinetics for energy minimization), this problem can be ameliorated in case of strong dependency of the CPM on continuous models, for which dynamics are well-defined, as shown in chapters 8 and 9.

### 3.2.3 Other cell-based models

The CPM framework is suitable to describe biological processes in which cell shape and cell motility are important determinants. However, many multicellular systems can be accurately described with cells as discrete interacting entities without explicitly accounting for cell shape or cell motility. Morpheus supports this type of modeling with a number of cell-based modeling formalisms.

Coupled ODE lattices, e.g. figure 3.1D1, are models in which each cell is spatially represented as a point-like object in a regular lattice, while its intracellular dynamics are governed by ODE models. These cells are coupled by the interactions between variables in directly adjacent cells. In cellular automata, e.g. figure 3.1E1, cells are represented as discrete lattice sites, but also have a discrete state space. Transitions among internal states are governed by rules that specify a cell's new state as a function of the current state and the state of the cells in the neighborhood. To coarsely represent cell motility in such models, one can include the asynchronous exchange of cell positions, upon some condition. In this way, it is possible to simulate interacting particle systems, as shown in figure 3.1E3. Even when cell motility is excluded, cell shape can be explicitly modeled by importing cell shapes from external images. To simulate processes on cellular geometries recorded from experimental data, Morpheus supports importing 2D and 3D TIFF images, e.g. figure 3.1D4 and E4.

### 3.2.4 Modularity and versatility

The basic components of the numerical simulation of the model formalisms described above depend on the availability of computational concepts for e.g. regular lattices, cell populations, diffusion solvers, cell motility and solvers for systems of tightly coupled equations. By encapsulating these basic components separately into modules, these modules can be combined in various ways. This allows a wide range of modeling



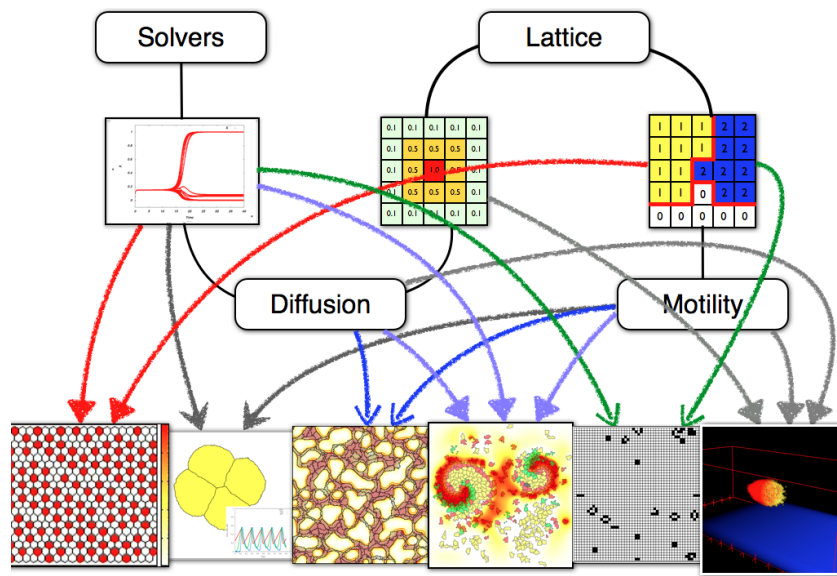


Figure 3.4.: **Modularity of modeling components.** Morpheus offers various modules: (1) solvers for (differential) equations, (2) lattices to spatial models, (3) diffusion solvers, and (4) cellular Potts models for cell motility. These modules can be combined in various ways to construct a rich variety of model formalisms as shown by the example in the bottom row.

formalisms to be constructed, in addition to the core formalisms, as illustrated in figure 3.4.

### 3.2.5 Computational efficiency

#### *Parallelization*

Morpheus uses multithreading to parallelize computations of differential equations on multicore (shared-memory) computers. Since the dynamics of intracellular systems of ordinary, stochastic or delay differential equations can be calculated independently for each cell for the periods between the exchange information between cells, these systems are computed in parallel, depending on the available number of cores. For reaction-diffusion systems, the reaction step is parallelized by decomposing the lattice along the y-axis and compute rows in parallel.

Although others have developed parallel implementations of the CPM using either message passing interface (MPI)<sup>116</sup> or graphical processing units (GPU)<sup>117</sup>, Morpheus does not support such parallelization. Instead, the computational load in simulating the CPM is reduced by avoiding calculation of updates that cannot change the configuration, i.e. updates where  $\sigma_{x_{\text{target}}} = \sigma_{x_{\text{source}}}$ . This is done by tracking the boundaries of cells, i.e. those lattice sites that can potentially change the configuration, and sampling from this set of lattice sites, instead of sampling all lattice sites. In this way, a large performance gain can be achieved (especially with large cell sizes) at the cost of a small memory overhead.

#### *Performance and scalability*

The performance and scalability of Morpheus simulations heavily depend on the type of (multi-scale) model that is being simulated. It is therefore difficult to make general statements on the computational efficiency. However, we can test the performance on a set of “benchmark” models that form the modules from which more complex model can be constructed. To test performance, we measured the execution time, memory usage and scalability for ODE lattices, reaction-diffusion (PDE) models, cellular Potts models (CPM) and a multiscale model (CPM+PDE), using the example models included in Morpheus. The results, presented in Appendix A.1 on page 159, show the execution time and memory consumption for these models as well as their scalability in terms of problem size and scalability in terms of efficiency of multi-threading.

## 3.3 GRAPHICAL USER INTERFACE

The growing complexity of computational models provide challenges to the management of the modeling work flow. Therefore, Morpheus has been designed with a focus on usability and work flow management. The graphical user interface (GUI), shown

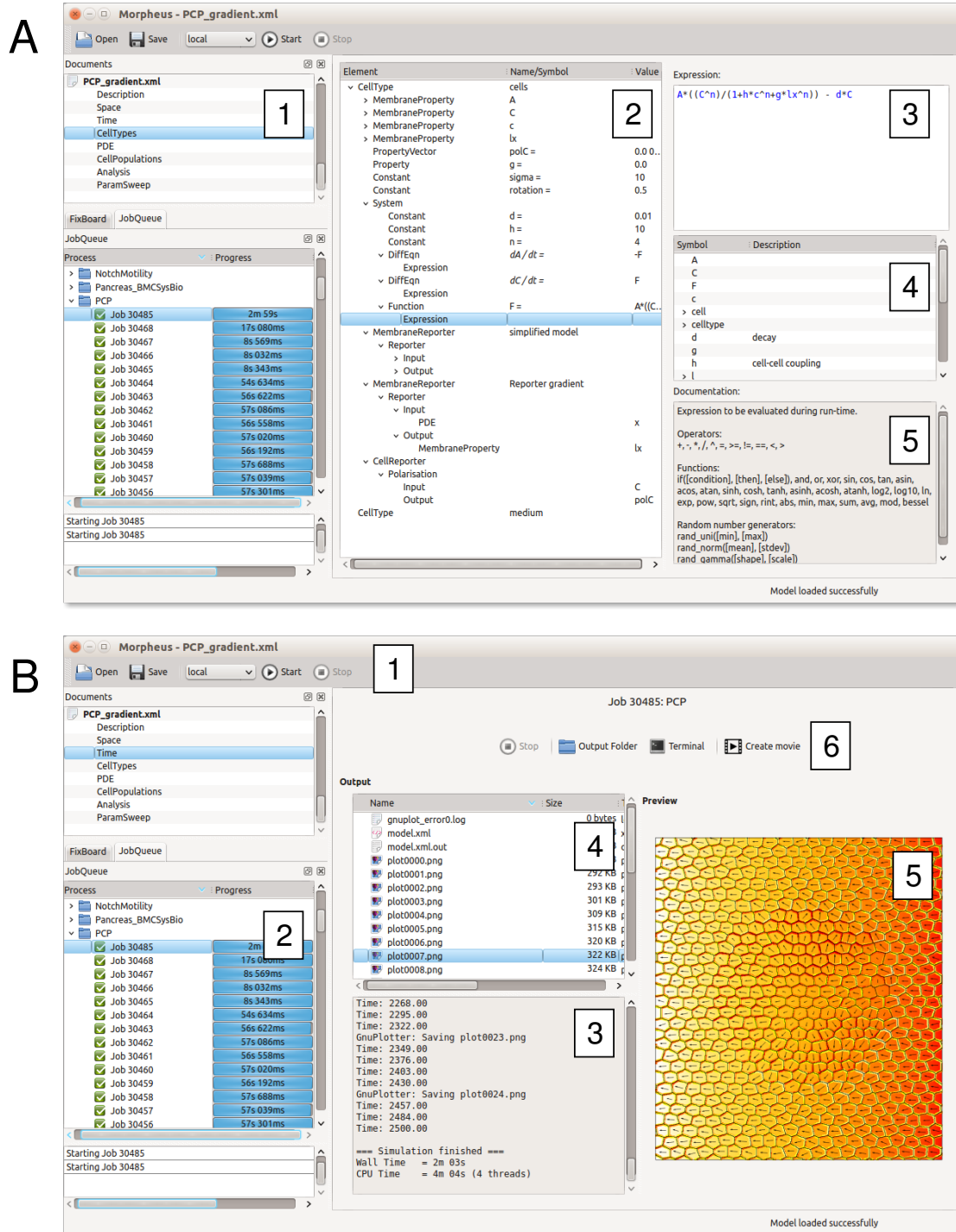


Figure 3.5.: **The graphical user interface provides tools for model editing, simulation and visualization.** (A) Model editor. (A1) Document view with main model elements. (A2) Overview of selected CellTypes element. (A3) Attribute and expression editor. (A4) Overview of available symbols. (A5) Context-sensitive documentation. (B) Job queue and result browser. (B1) Toolbar to start/stop simulation. (B2) Upon starting a job, simulations are added to the job queue. (B3) Standard output is displayed and updated during simulation. (B4) Simulation results are saved to files. (B5) Text and image files can be visualized using the preview panel. (B6) Output files can be opened in external terminal or file browser for postprocessing, and movies can be generated from images.

in figure 3.5, is an important part of the usability, providing an intuitive interface to construct models and execute simulations. In addition, it includes a number of tools to browse and visualize simulation results and to perform parameter exploration. The GUI is implemented in C++ using the platform-independent Qt application framework.

The main functionality of the GUI is to construct and edit models in MorpheusML format. Model elements and formalisms are added and removed by editing the tree-like structure (fig. 3.5A2) whereas parameters and mathematical expressions are specified in the attributes panel (fig. 3.5A3). Multiple models can be opened simultaneously to allow comparison of results as well as copy/pasting elements between models. Models, constructed in the GUI, are automatically written to file in MorpheusML format upon saving or when executing a simulation. The GUI provides a scheduling system that allows multiple simulations to be executed in sequentially or in parallel, depending on the available number of processors. In addition to a standard execution mode in which output is written to file, the GUI also supports an interactive execution mode in which all visual output is directed to an on-screen terminal. The job archive provides an overview of the job status (i.e. pending, running, done, error) (fig. 3.5B2). Using an SQL database backend, it enables the user to browse, sort and visualize simulation results as well as generating animations (fig. 3.5B4-6). It also restores (old) simulation models and parameters sweep from the archive.

Batch processing for parameter exploration or sensitivity analysis is supported by the ParamSweep tool. This allows all model parameters to be selected for batch processing and given user-specific or generated sequences of values in linear or logarithmic intervals. Multiple parameters can be explored combinatorially or can be paired to vary in synchrony. Remote execution enables users to transparently execute (batch) simulations on high performance computing resources (HPC). In this case, models are transferred to the HPC using secure file transfer (sftp) and submitted to a batch system (LSF or SLURM) for job scheduling. Results can be transferred back to the local computer after simulation or synchronized during its execution.

The GUI drastically reduces the learning curve involved in computational modeling and allows users to focus on the biological and mathematical aspects of modeling, instead of the computational ones. Experience using Morpheus in education has shown that students, with both mathematical and biological backgrounds, are able to construct and execute simulation models within hours.

### 3.4 MORPHEUSML

To facilitate the construction of a versatile array of modeling formalisms (table 3.1) from within user-friendly graphical interfaces (figure 3.5), we have designed a novel model description language, MorpheusML. Similar to a markup languages such as SBML<sup>93</sup>, this is a declarative XML-based format to represent complex dynamical models using domain-specific biological and mathematical terminology. However, whereas SBML limits itself to the representation of biochemical pathways, MorpheusML is specifically designed to represent models of spatial and multicellular systems, including multiscale models.

Models in MorpheusML format contain a complete specification of the simulation. This includes the definition, configuration and parameterization of (sub)models as well as the specification how these (sub)models are interlinked. It also includes details on the numerical simulations such as the simulation time, spatial discretization, initial conditions and the configuration of visualization and data output. During simulation, the complete state of a simulation can be stored in the same file format. The encapsulation of complete model description in a single XML file, render them suitable for archiving as well as model exchange between users.

#### 3.4.1 Declarative and domain-specific

The declarative nature of MorpheusML serves to separate modeling from implementation. That is, models in MorpheusML describe *what* processes are to be simulated rather than *how* this should be accomplished. This distinguishes declarative languages from imperative programming language such as C++ or Python that focus on the description of algorithmic control flow. This also enables a separation of concern between modelers and programmers, where the former specify what process should be simulated and the latter specify how this should be accomplished numerically.

Moreover, the use of biological and mathematical terminology, instead of programming constructs, makes for a more natural descriptions of models of complex biological processes. MorpheusML is composed of human-readable tags to represent the components of biological processes as well as mathematical constructs to define their dynamics and relations.

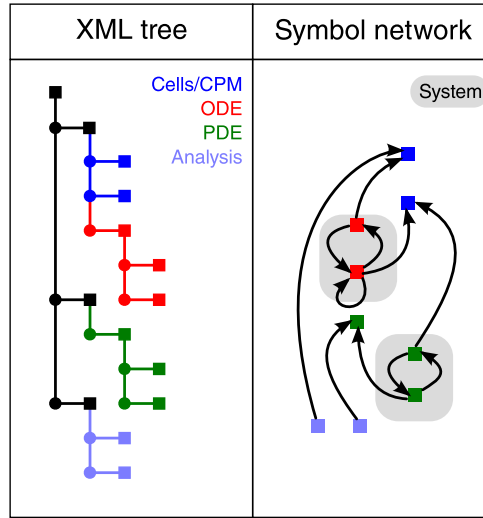


Figure 3.6.: **Two-tiered architecture of MorpheusML.** Left: The XML hierarchy provides an easily parsable format to represent the structure of simulation model and its sub-models. Right: The usage of symbolic identifiers and mathematical expressions allows the representation of coupling between submodels into networks of interdependencies. The combination of these two tiers of description enables the representation of arbitrarily complex multiscale models.

### 3.4.2 Two-tiered architecture

MorpheusML represent complex multiscale models of multicellular systems in a two-tiered architecture, as depicted in figure 3.6. On the one hand, it represents models uses a convenient hierarchical tree structure. On the other hand, symbolic references between different parts of this structure enable the representation of complex network of interactions.

Representation of data in eXtensible Markup Language (XML) has a number of key advantages: (1) it stores information in a well-structured fashion that can be easily parsed and validated, (2) it allows human-readable domain-specific terminology and (3) it can be extended in a straightforward fashion. The XML represents this information in a hierarchical tree-like structure that reflects the structure of the modeled biological system. For instance, the main element `CellPopulations` can contain a `Population` that contains multiple `Cells`, each of which may contain multiple lattice `Nodes`. Similarly, intracellular dynamics are modeled using a `Systems of DiffEqn` within a `CellType`, while the PDE describing extracellular dynamics is defined in its own element outside of `CellTypes`. The XML structure is convenient to represent the hierarchy between the components of a model.

However, the hierarchical structure of the XML format is ill-suited to describe the network of interactions and feedbacks between these model components. To describe

these interactions, MorpheusML uses symbolic identifiers. Symbolic identifiers and references establish interactions and feedbacks between (sub)models to represent the network-like complexity in biological processes. Symbolic identifiers, or symbols for short, can be specified to represent user-defined model variables such as cell-bound properties (Property) or concentrations of species in a reaction-diffusion model (Layer) and can also be specified for simulation-related constants and variables such as lattice size and current time of simulation.

Symbols can be used in mathematical constructs and expressions to define relations between model components. Using a so-called “white-box approach” (as opposed to black-box approach), all symbols of a modeling component are available as potential coupling points with other components<sup>118</sup>. This provides a convenient way to integrate different (sub)models by defining symbolic identifiers in one (sub)model and using them in another (sub)model.

The combination of a hierarchical XML structure on one hand and networks of symbolic identifiers on the other hand provides a powerful way to describe complex multiscale models in a declarative fashion.

### 3.4.3 Mathematical constructs

MorpheusML includes a number of mathematical constructs to specify algebraic expressions and define relations among model components (table 3.2). Local and global constants (Constant and Global) as well as cell-bound and PDE variables (Property and Layer) associate user-specified symbolic identifiers with values. Additionally, specific containers are available for (x,y,z) vector variables (PropertyVector) and variables for use in delay (differential) equations (DelayProperty). Mathematical expressions can be specified as Functions, Equations, Rules or DiffEqns (differential equation) in which the right-hand-side expression can be given in conventional infix notation.

Sets of tightly coupled (differential) equations can be defined in Systems which provides an environment for synchronously updated equations and are associated with a user-specified integration timestep and numerical solver (e.g. Runge-Kutta). Similarly, an Event provides an environment to trigger a set of synchronously updated expressions, based on a user-specified Condition, allowing time-continuous event-based simulations.

With these mathematical constructs, MorpheusML offers a layer of flexibility to the description of complex dynamic models within a fully declarative language. Moreover, it enables the direct conversion of other model formats into MorpheusML. In particular, owing to the high degree of similarity, models of biochemical networks in Systems Biology Markup Language (SBML) format can be translated into MorpheusML. The graphical user interface provides a tool to automatically convert SBML models into MorpheusML models.

```

<?xml version='1.0' encoding='UTF-8'?>
<MorpheusModel version="1">
  <Description>
    <Title>LateralSignaling</Title>
    <Details>W de Back, JX Zhou, L Bruschi
      Roy. Soc. Interface 10(79): 20120766, 2012.
      http://dx.doi.org/10.1098/rsif.2012.0766
    </Details>
  </Description>
  <Space>
    <Lattice class="hexagonal">
      <Size value="20 20 0"/>
      <BoundaryConditions>
        <Condition boundary="x" type="periodic"/>
        <Condition boundary="y" type="periodic"/>
      </BoundaryConditions>
    </Lattice>
  </Space>
  <Time>
    <StartTime value="0"/>
    <StopTime value="30"/>
    <TimeSymbol symbol="t"/>
  </Time>
  <CellTypes>
    <CellType class="biological" name="cells">
      <Property symbol="X" value="0.0" name="Ngn3"/>
      <Property symbol="Xn" value="0.0" name="Ngn3-Neighbors"/>
      <Property symbol="Y" value="0" name="Ptfla"/>
      <Property symbol="Yn" value="0" name="Ptfla-neighbors"/>
      <NeighborsReporter mapping="average">
        <Input symbol-ref="X"/>
        <Output symbol-ref="Xn"/>
      </NeighborsReporter>
      <NeighborsReporter mapping="average">
        <Input symbol-ref="Y"/>
        <Output symbol-ref="Yn"/>
      </NeighborsReporter>
      <System solver="heun" time-step="0.02">
        <Constant symbol="a" value="1"/>
        <Constant symbol="b" value="21"/>
        <Constant symbol="c" value="1"/>
        <Constant symbol="n" value="4"/>
        <Constant symbol="th" value="1e-4"/>
        <Constant symbol="noise" value="1e-4"/>
        <DiffEqn symbol-ref="X">
          <Expression>((th / (th + a*Xn^n)) - X)
            + rand_norm(0.0,noise)
          </Expression>
        </DiffEqn>
        <DiffEqn symbol-ref="Y">
          <Expression>((th + b*(Y * Yn)^n) / (th + c*X^n + b*(Y * Yn)^n)) - Y
            + rand_norm(0.0,noise)
          </Expression>
        </DiffEqn>
      </System>
    </CellType>
  </CellTypes>
  <CellPopulations>
    <Population size="0" type="cells">
      <InitCA/>
    </Population>
  </CellPopulations>
  <Analysis>
    <Gnuplotter clean="true" interval="2.5">
      <Terminal name="png"/>
      <Cells symbol-ref="Y" min="0" max="1.0"/>
    </Gnuplotter>
  </Analysis>
</MorpheusModel>

```

Annotation

Space and Time

CellType

ODE

Initialization

Visualization

Figure 3.7.: Listing of lateral signaling model in MorpheusML format. (Caption on next page.)



Table 3.2.: Mathematical constructs in MorpheusML

Element	Description
<b>Containers</b>	
Constant	Constant value of type double with local scope, i.e. valid within the <code>CellType</code> or <code>System</code> it is defined in.
Global	Variable value of type double with global scope.
Property PropertyVector DelayProperty	Cell-bound variable. <code>Property</code> and <code>DelayProperty</code> are of type double. <code>DelayProperty</code> has attribute <code>delay</code> to set the lag between assignment and return of value. <code>PropertyVector</code> defines Euclidean vector in space delimited format “x y z”.
Layer	PDE model variable, i.e. species in reaction-diffusion system. Diffusivity of a <code>Layer</code> is specified in attribute <code>diffusion</code> .
<b>Expressions</b>	
Function	Mathematical expression. Computes a value (double) for the output symbol it defines, but does not assign it to a variable. Updated whenever when output symbol is referenced. May not contain algebraic loop.
Equation	Mathematical expression. Computes a value (double) and assigns it to the variable it references. Updates are scheduled depending on its symbol dependencies. May not contain algebraic loop.
Rule	Mathematical expression that defines a (recurrence) equation for use in environments such as <code>System</code> and <code>Event</code> . Scheduled according to <code>System/time-step</code> . May contain algebraic loop and self-references.
DiffEqn	Mathematical expression that defines a differential equation. Only allowed in <code>System</code> environment. May contain algebraic loop and self-references.
<b>Reporters</b>	
Reporter NeighborsReporter PDEReporter ...	Explicit data mappings. Computes a statistic (average, mean, etc.) of the input data and assigns this to the output symbol. Updates are scheduled depending on its symbol dependencies.
<b>Environments</b>	
System	Environment for tightly coupled sets of differential equations and rules that are synchronously updated. Scheduled according to user-specified <code>System time-step</code> and <code>time-scaling</code> .
Event	Environment for conditional or timed events. Events are triggered (1) periodically (2) whenever <code>Condition</code> holds true or (3) whenever <code>Condition</code> changes from false to true (for SBML compatibility). Updates are scheduled according to <code>time-step</code> if specified or depending on its symbol dependencies otherwise.

Figure 3.7.: **Listing of lateral signaling model in MorpheusML format.** This model specifies a coupled ODE lattice with lateral inhibition and lateral stabilization, as developed in chapter 5. After annotation notes, the size and structure of the lattice are defined as well as the simulation time. Then, the cell type is defined with cell properties, neighborhood dependencies and the systems of differential equations. Finally, the population is initialized and the visualization output is specified.

#### 3.4.4 Encapsulation

Models in MorpheusML format provide a complete specification of a simulation model. Figures 3.7 and 3.8 show full listings of two models that are developed in respectively chapters 5 and 9 of this thesis. Despite the fact that they provide a complete model specification, models in this format are concise and, moreover, human readable.

Encapsulation of in a single XML file has various advantages, as compared to models that are represented in programming code. For instance, it allows executed simulation models to be archived together their results and be restored later. This enables users to keep track of different variations of a model. It also facilitates the exchange of models among users. MorpheusML model files (typically  $\approx 3\text{Kb}$ ) can be easily sent over email and started by a collaborator by drag/drop, without need for configuration or compilation. Thus, MorpheusML opens up the possibility for the creation of an online public model repositories for multiscale multicellular simulation models analogous to the BioModels repository\* for models in SBML format.

#### 3.4.5 Semantic model analysis

The markup language enables Morpheus to perform a semantic analysis of models to identify logical errors that would not be identified by compilers. For instance, whereas general-purpose compilers would not raise errors in case of a double assignment (in which the second assignment overrides the first), these are detected Morpheus as errors that invalidate a simulation model. Similarly, algebraic loops (in which e.g.  $a$  depends on  $b$ , while  $b$  depends on  $a$ ) are detected in Morpheus as unsolvable constructs (unless they appear in a synchronously updated System).

### 3.5 AUTOMATED MODEL INTEGRATION

Multi-scale modeling requires the integration of time-discrete cell-based models with time-continuous models for intra- and extracellular dynamics. In Morpheus, cell-based models can be linked to ordinary, stochastic or delay differential equations as well as

---

\* <http://www.ebi.ac.uk/biomodels-main>

```

<?xml version='1.0' encoding='UTF-8'?>
<MorpheusModel version="1">
  <Description>A Köhn-Luque, W de Back, Y Yamaguchi, K Yoshimura, M A Herrero, T Miura
    Physical Biology, 10(6):066007, 2013
    doi:10.1088/1478-3975/10/6/066007</Details>
  </Description>
  <Space>
    <Lattice class="square">
      <Size value="200 200 0"/>
      <BoundaryConditions>
        <Condition boundary="x" type="periodic"/>
        <Condition boundary="y" type="periodic"/>
      </BoundaryConditions>
      <NodeLength unit="micron" value="2"/>
      <Neighborhood>
        <Order>2</Order>
      </Neighborhood>
    </Lattice>
  </Space>
  <Time>
    <StartTime value="0"/>
    <StopTime value="4000"/>
  </Time>
  <CellTypes>
    <CellType class="biological" name="Angioblasts">
      <VolumeConstraint>
        <Strength value="1"/>
        <Target value="90"/>
      </VolumeConstraint>
      <Chemotaxis>
        <Layer symbol-ref="w"/>
        <Strength symbol-ref="s"/>
      </Chemotaxis>
      <Property symbol="s" value="3e7" name="chemotactic strength"/>
      <Property symbol="cell" value="1.0" name="cell"/>
    </CellType>
    <CellType class="medium" name="medium">
      <Property symbol="cell" value="0.0" name="cell"/>
    </CellType>
  </CellTypes>
  <CPM>
    <Interaction>
      <Contact type1="medium" type2="Angioblasts" value="1.6"/>
      <Contact type1="Angioblasts" type2="Angioblasts" value="3.2"/>
    </Interaction>
    <MetropolisKinetics temperature="1" stepper="edgelist" />
    <MCSDuration value="1.0"/>
  </CPM>
  <PDE>
    <Layer symbol="u" name="VEGF_s">
      <Diffusion rate="58.7" unit="μm²/s"/>
      <Initial>
        <Expression>1.5e-6</Expression>
      </Initial>
    </Layer>
    <Layer symbol="v" name="Free ECM">
      <Diffusion rate="0.001" unit="μm²/s"/>
    </Layer>
    <Layer symbol="w" name="VEGF_b">
      <Diffusion rate="0" unit="μm²/s"/>
    </Layer>
    <System solver="heun" time-step="1.0">
      <Constant symbol="gamma" value="5e-3" name="Production ECM"/>
      <Constant symbol="k_on" value="8.5e-4" name="Binding rate VEGF/ECM"/>
      <Constant symbol="k_off" value="3.6e-3" name="Unbinding rate VEGF/ECM"/>
      <Constant symbol="delta" value="2.6e-6" name="Decay VEGF"/>
      <DiffEqn symbol-ref="u">
        <Expression>- k_on*u*v+k_off*w-delta*u</Expression>
      </DiffEqn>
      <DiffEqn symbol-ref="v">
        <Expression>gamma*cell - k_on*u*v+k_off*w</Expression>
      </DiffEqn>
      <DiffEqn symbol-ref="w">
        <Expression>k_on*u*v-k_off*w</Expression>
      </DiffEqn>
    </System>
  </PDE>
  <CellPopulations>
    <Population size="0" type="Angioblasts">
      <InitRectangle cells="200" type="regular">
        <Dimensions size="200 200 0" origin="0 0 0"/>
      </InitRectangle>
    </Population>
  </CellPopulations>
</MorpheusModel>

```

Annotation

Space and Time

CellTypes

CPM

PDE

Initialization

Figure 3.8.: Listing of vascular morphogenesis model in MorpheusML format. (Caption on next page.)

Figure 3.8.: This model defines a cellular Potts model, coupled to a reaction diffusion system, as developed in chapter 9. First, annotation notes and the definition of the lattice and simulation duration are given. In `CellTypes`, the cell behavior is specified including volume conservation and chemotactic behavior. The details of the CPM, such as the interaction energies and temperature, are specified. In the PDE element, the reaction-diffusion system is defined, including the measured kinetic parameters. Finally, the initial configuration of the population is specified. For brevity, the specification of analysis methods are not listed here.

to reaction--diffusion (PDE) models. MorpheusML facilitates this through the specification of links between (sub)models with the help of symbolic identifiers.

For the user, a link between sub-models is established by defining a symbol in one submodel and using it as an input in another sub-model, providing a convenient way to construct and explore complex multiscale biological systems using integrative models. During simulation, Morpheus makes the data accessible between sub-models and, if necessary, mapping or transforming it to make it suitable for the target submodel. Moreover, updates of the various submodels are appropriately scheduled, by determining the correct order and the frequency of updates, as to guarantee that up-to-date data is used in all computations. Both tasks are handled automatically as far as possible, based on user-specified time-steps and symbolic interdependencies.

#### *Integration of spatial models*

Integration of spatial model formalisms, i.e. cell-based and reaction--diffusion models, requires that the data from one submodel is accessible to the other sub-model. In Morpheus, data is not copied between submodels, but is directly accessible through symbolic identifiers. Yet, the data must be accessed in a way that is appropriate for the model that uses the symbol. The model that uses a symbol determines the lattice sites for which the symbol is resolved, whereas the model that defined it determines the value of the symbol at those lattice sites.

Using the convention that spatial discretization is equal for all spatial models, the mapping is often trivial enough to be handled automatically. In other cases, if a sum or average over a lattice domain is required, this mapping must be made explicit by the user using a `Reporter`.

#### *Scheduling updates*

Integration of time-continuous ODE and PDE models with time-discrete CPM models and various auxiliary mathematical constructs requires a careful scheduling of numerical updates. While the general simulation schedule is executed in fixed order, some processes must be scheduled according to the symbolic interdependencies in order to

guarantee correctness of simulation results. This affects both the order of executing updates and the frequencies at which they are updated.

The order of execution is independent of the order in which model components are specified in the model description file. Rather, updates of temporal processes (CPM, System, Diffusion) are scheduled according to the fixed schedule and updates of sequential processes such as Equations and Reporters are scheduled according to their symbolic interdependencies. These are ordered according to the dependencies in their input and output symbol based on the rule:

- Before updating a process, all its input symbols must be updated.

This is achieved by scheduling all processes that have these symbols as an output prior. Note that this is only possible if no algebraic loops or circular dependencies exist between these processes. Therefore, such loops are only allowed within the System environment.

The update intervals of sequential processes such as Equations, Reporters and Events are automatically determined by propagation of the intervals of their input and output symbols, according to the following rules:

- The process is updated as often as its output symbol(s) are used.
- The process is not updated more often than its input symbol(s) can change.

The former ensures correctness in that up-to-date data is used in all processes, while the latter optimizes computational performance by preventing redundant computations. Note, for instance, that this implies that a process is not scheduled and computed if their output symbol is not used.

The full simulation schedule, including the order and intervals of all model elements, is calculated at the end of initialization and displayed in the output text box of results view (fig. 3.5B3) of the graphical user interface.

### 3.6 SOFTWARE ARCHITECTURE AND EXTENSIBILITY

The software itself consists of two stand-alone applications, the graphical user interface (GUI) and the simulator. These applications communicate through the exchange of XML and XSD files. During compilation, an XML schema description (XSD) file is constructed that describes the contents as well as the rules and constraints of valid model description files in MorpheusML that the simulator expects as an input. This information is used by the GUI to assist the user in the construction and editing of valid models. Once a simulation is executed, the GUI generates a XML file, based on the user input and starts the simulator with the XML file as an input argument.

This separation of GUI and simulator has several advantages. It allows Morpheus to be used in a headless fashion without a graphical interface. Moreover, it implies that

the GUI and simulator do not need to run on the same computer. Therefore, it is also possible to construct models with the GUI on a local desktop machine while executing (parallel) simulations on a remote high performance computer by sending XML files over a network connection.

It also enables Morpheus to be used at two different levels. Morpheus can be used as an off-the-shelf application that is shipped as binary packages for all major operating systems (Linux, Mac OSX and MS Windows). However, it is also a software framework that can be extended through its plugin architecture, although this requires building the code from source code. As a binary application it is aimed to be easy-to-use and without any requirements for experience with computational techniques such as compilation of source code. Yet, as a software platform it provides the transparency and customizability similar to a software library.

### 3.6.1 Extensibility

Morpheus' plugin architecture allows customization and extensibility. Plugins are typically small pieces of code providing a specific feature. These are isolated from the core application and can be built independently from each other. The plugin architecture enables third-party users and developers to create custom features for modeling, initialization or analysis. In fact, most standard features of Morpheus are implemented as plugin themselves.

Morpheus defines interfaces for various types of plugins that differ in the order and time they are executed. Each type has its custom interface with functions that the derived class, the plugin, is required to implement. Initializer plugins set up the initial spatial configuration of a population of cells or the concentrations in reaction-diffusion systems and are executed during initialization of the simulation. For instance, `InitRectangle` is a plugin that initializes a population of cells in evenly spaced over a rectangular (or cubic) area and `TIFFReader` initializes cells according to a 2D or 3D image in TIFF format. `CellType_MCS_Listener` or `TimeStepListener` plugins are, on the other hand, executed during simulation either at each Monte Carlo step or at the time step set for the numerical solvers for continuous models. The `Proliferation` plugin, for example, is executed every Monte Carlo step and models cell division depending on a certain user-specified condition while the `NeighborsReporter` is executed at every time step to calculate and report a weighted average of the values of a cell property of adjacent cells. Plugins for run-time data analysis or visualization, such as the versatile `Gnuplotter`, use the `Analysis_Listener` interface and are called at user-specified intervals.

Plugins are written in C++ and consist of a header file (`myplugin.h`) and an implementation or source file (`myplugin.cpp`) (see figure 3.9). The header file provides a forward declaration of a class and declares public and private variables and functions. In addition, it defines the type of plugin interface it inherits. The source file contains

## Header

## myPlugin.h

```

#include "core/interfaces.h"
#include "core/simulation.h"
#include "core/celltype.h"

class MyPlugin : public Analysis_Listener
{
public:
    DECLARE_PLUGIN("MyPlugin");
    NetworkLogger() {};
    virtual void loadFromXML(const XMLNode);
    virtual void notify(double time);
    virtual void init(double time);

private:
    string symbolstr, celltypestr;
    SymbolAccessor<double> symbol;
    shared_ptr<const CellType> celltype;
};

```

Inherit interface

Override functions

Symbol references

## Implementation

## myPlugin.cpp

```

#include "myPlugin.h"
REGISTER_PLUGIN(MyPlugin);

void MyPlugin::loadFromXML(const XMLNode Node)
{
    Analysis_Listener::loadFromXML( Node );
    getXMLAttribute(Node, "celltype", celltypestr);
    getXMLAttribute(Node, "symbol", symbolstr);
}

void MyPlugin::init(double time)
{
    Analysis_Listener::init(time);
    celltype = CPM::findCellType(celltypestr);
    symbol = SIM::findSymbol<double>(symbolstr, celltype);
}

void MyPlugin::notify(double time)
{
    Analysis_Listener::notify(time);
    for(uint c=0; c < celltype->getCellIDs().size(); c++){
        double value = CPM::getCell( cells[c] ).get( symbol );
        // do something
    }
}

```

Read parameters

Initialize references

Use symbols

## XSD

## myPlugin.xsd

```

<xs:schema>
  <xs:complexType name="MyPlugin">
    <xs:annotation>
      <xs:documentation>Documentation of MyPlugin
    </xs:documentation>
    </xs:annotation>
    <xs:attribute name="interval" type="cpmDouble" use="required"/>
    <xs:attribute name="celltype" type="cpmCellTypeRef" use="optional"/>
    <xs:attribute name="symbol" type="cpmSymbolRef" use="required"/>
  </xs:complexType>
</xs:schema>

```

Documentation

Attributes

Figure 3.9.: **Components of a plugin.** A header file declares the plugin class and inherited interface functions (here from the `Analysis_Listener` interface). An implementation file reads parameters from XML file, provides the methods of the plugin and overrides the interface functions. An XSD file defines the rules and constraints for valid input XML to configure the plugin and provides documentation about the plugin.

the implementation of the methods of the plugin and must implement the interface functions. Plugins are configured and parameterized by XML and should therefore implement a `loadXML()` function that reads the XML configuration for the plugin. The structure and content of the XML that the plugin expects as input must be defined in a XML description schema (XSD) file (`myplugin.xsd`) (see figure 3.9). This schema is used by the graphical interface to ensure correct usage of the plugin and to provide documentation on its functionality.

### 3.7 CONCLUSIONS

Morpheus is a versatile modeling environment for the simulation of multiscale and multicellular systems, designed with a focus on usability. It separates the tasks of modeling from implementation and thereby enables a division of labor between biological modelers, who describe *what* should be modeled, and computational biologists, who implement *how* this should be accomplished numerically. For the former, Morpheus provides intuitive graphical interfaces for model construction and execution. For the latter, it provides a C++-based extensible framework with a plugin architecture.

Morpheus makes an important step towards making dry computational modeling accessible to wet lab biologists. With its focus on usability, it provides a tool that can integrate modeling into the common work flow of biologists, along with commonly used software for statistical analysis or image analysis.

So far, the software has been used in a number of published research studies ranging from cell fate decisions, spatial patterning<sup>119</sup> and morphogenesis<sup>102,115</sup> as well as a number of unpublished work including studies on liver tissue architecture, regeneration in planaria and planar cell polarity. It has also been used in a number of courses teaching computational modeling to biologists, including a course for the DIGS-BB graduate programme and a workshop for the German Stem Cell Network.

Despite the versatility of modeling formalisms that can be constructed in Morpheus, these are subjected to several important limitations. For instance, whereas irregular lattices are known to be more robust against anisotropic lattice artefacts, only regular lattices are currently supported. The use of finite difference solvers with fixed user-defined time-steps can be numerically inefficient and precludes simulation of stiff systems. Morpheus does not support the simulation of reaction-diffusion-advection systems. And currently, cell motility can only be modeled using the (rather detailed) cellular Potts model or the (rather coarse-grained) interacting particle systems. Improvements on all these levels remain for future development.

Perhaps the key novelty in Morpheus' design is the establishment of MorpheusML as a novel markup language for models of multicellular systems and the implementation of an interpreter for this language that facilitates model integration. Although MorpheusML is designed for use within Morpheus alone, the concepts underlying the structure of the language and its interpretation may prove valuable for the future estab-



lishment of an SBML-like standard exchange format to share models of multicellular systems between different software platforms<sup>94</sup>.

The continuation and further development of Morpheus, including MorpheusML, crucially depend on the involvement of other researchers and contributions from external software developers. We therefore aim to transform Morpheus development into an open source project. For the upcoming version, next to the binary applications, all source code will be released under open source license.



## Part II

### CELL FATE REGULATION AND PATTERNING

---

## INTRODUCTION

---

The computational methods described in Part I enable the investigation of the development of multicellular systems as an interplay between cell fate specification, patterning and morphogenesis. Here, a case study is presented in which these methods are used to study how the pattern formation within the tissue can affect the fates of its constituent cells. This is done by spatially coupling discrete cells in a tissue model to their intracellular models of gene regulatory networks. Specifically, we investigate the dynamics of specification and spatial patterning of endocrine and exocrine cells in the pancreas, during embryonic development as well as during reprogramming *in vitro*.

### 4.1 CELLULAR PLASTICITY AND REPROGRAMMING

During embryonic development, pluripotent cells progressively become more differentiated to obtain a final functional state. Conrad Hal Waddington compared this process to a marble rolling down and slope becoming shunted into one of the several valleys (figure 4.1)<sup>120</sup>. As the cell reaches a terminally differentiated state, the hills around the valley are steep enough to irreversibly lock the cell in this state. This picture, inspired by dynamical systems theory, remains one of the most powerful metaphors of cell fate regulation<sup>23</sup>.

Yet, due to recent advances in stem cell biology and cellular reprogramming, we now know that fully differentiated cells retain a level of plasticity. Given the right stimuli, adult cells can be reverted back to a pluripotent state and even converted from one cell type to another without passing through a state of pluripotency (figure 4.1)<sup>121</sup>.

These technologies hold their many promises for basic research and as well as for regenerative medicine, since reprogrammed cells could replace dysfunctional or depleted cells in degenerative diseases. For diabetes mellitus, for instance, cells that are converted from other tissues could be replacing the dysfunctional insulin-producing  $\beta$ -cells of the pancreas<sup>122,123</sup>. Among the many attempts to convert cells from various source tissues to new  $\beta$ -cells<sup>124</sup>, in a recent landmark study, Douglas Melton and colleagues have demonstrated the possibility of *in vivo* reprogramming of acinar pancreatic cells to functional  $\beta$ -cells in mice by overexpression of only three key transcription factors: Ngn3, Pdx1 and MafA<sup>125</sup>.

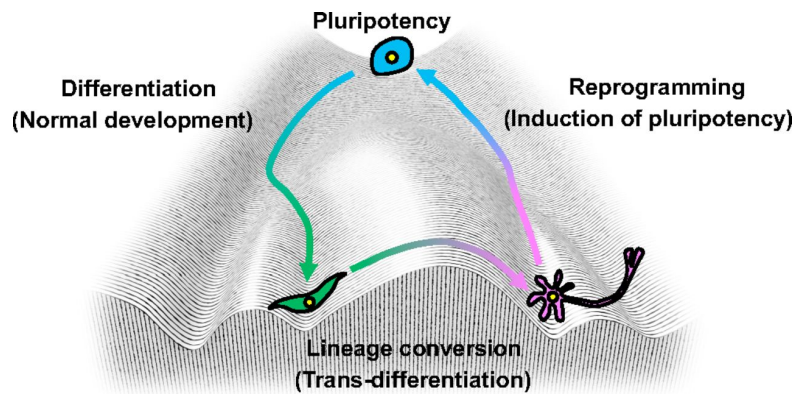


Figure 4.1.: **Cell fate changes on Waddington's epigenetic landscape.** Pluripotent stem cell can be induced from adult differentiated cells (reprogramming) and differentiated back to another cell lineage, mimicking normal development. Cells can also be converted to another lineage without passing through a state of pluripotency (trans-differentiation)<sup>121</sup>.

In these types of studies, the choice of source cell type and the combination of transcription factors to overexpress is based on (1) knowledge of the genes that code for key lineage-determining transcription factors that are activated during development and (2) the piecemeal data on the regulatory network that these transcription factors form. In most cases, viral vectors are constructed with different combinations of a selection of these genes with which source cells are infected *in vitro* or *in vivo*. Although the results of some of these studies are indeed impressive and promising, in a sense, such trial-and-error experiments are picking low hanging fruit and many hurdles remain before these reprogramming techniques are efficient and safe enough for human treatment.

Apart from many technical challenges, two pertinent problems remain. First, little is understood about the transcriptional dynamics that is involved in the transition from one cell state to the other. This is dangerous since it is plausible that cells may get trapped in unnatural states, not normally visited during development, possibly resulting in tumorigenesis<sup>126</sup>. Second, the commonly used technique of viral gene transfer is based on random insertion of the transgenes within the genome and is associated with high risks of tumorigenesis that prevent any clinical use<sup>127</sup>. Therefore, alternative approaches are actively being explored. These include the use of methods that avoid integration into the host genome such as episomal vectors, small molecules, RNA<sup>128</sup>, CRISPR/Cas9-based gene editing<sup>129</sup>. Additionally, there is increasing awareness of the facilitating roles that cell-cell communication plays during reprogramming and the maintenance of pluripotency, such as gap junctional communication<sup>130</sup> and cadherin-mediated signaling<sup>131</sup>. However, designing effective reprogramming protocols ultimately depends on a deep understanding of the regulation of cell fate stability,

for which tools from dynamical systems theory and computational modeling are indispensable.

#### 4.2 MATHEMATICAL MODELING OF TRANSCRIPTIONAL REGULATION

This has sparked renewed interest in theoretical understanding of the regulation of cell fate decisions<sup>23,132–139</sup> building on the early works of Waddington's epigenetic landscape<sup>120</sup>, Delbrück's notion of bistability<sup>140</sup>, Jacob and Monod's discovery of transcriptional regulation<sup>141</sup> and Kauffman's random genetic networks<sup>142</sup>. By adopting the framework of dynamical systems theory, cellular characteristics such as cell types, pluripotency and cell fate decisions can be understood as attractors, metastable states and trajectories in a high-dimensional state space, constrained by the regulatory network encoded in the genome. Moreover, it replaces the view of genetic circuits as causal pathways by the concept of network dynamics which enables a new understanding of cell fate decisions in development and can be used to design cellular reprogramming protocols<sup>143</sup>.

As a simple example of mathematical modeling of a transcriptional network, consider a small gene circuit of two mutually inhibiting genes  $x$  and  $y$  (see figure 4.2A). This system, known as a genetic toggle switch<sup>144</sup>, has the possibility to settle in either of two expression states. The change in expression of the genes can be represented using two coupled ordinary differential equations in which the first term models synthesis, repressed by the other gene, according to Hill kinetics and the second term represents degradation using mass action kinetics (fig. 4.2B). From these equations, the steady states and their stability can be calculated and the quasi-potential surface can be constructed with hills and valleys (fig. 4.2C,D), analogous to Waddington's epigenetic landscape. In this case, three steady states can be observed, two of which are stable, representing the decided differentiated fates, while the other is unstable, representing the undecided progenitor state. Assuming an initial condition in which both genes are not expressed, this cross-inhibition between these genes results in a binary cell fate decision.

This two-gene example may seem too simplistic in the context of the number of genes involved in regulating embryonic development. Yet, there is in fact a growing evidence that cross-inhibition between two key transcriptional factors governs binary cell fate decisions. Mutual inhibition between *Cdx2* and *Oct4*, for instance, controls the decision between the trophoectoderm and the inner cell mass fates in pluripotent embryonic cells and inhibition between *GATA1* and *PU.1* governs the decision between the erythroid and myeloid lineages in haematopoiesis<sup>145</sup>. Therefore, even such simple mathematical models can give insight into the nature of cell types, as stable or unstable attractors in the space of expression states, as well as the dynamics of regulation in cell fate decisions that is usually inaccessible to experimental observation.

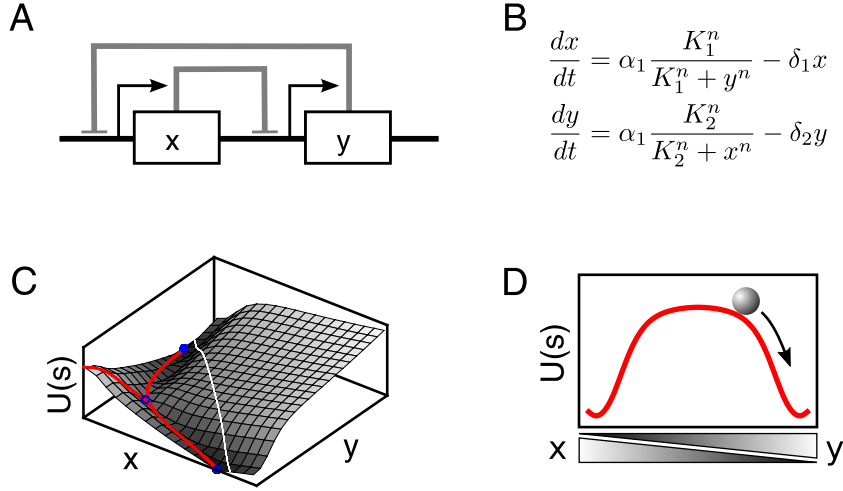


Figure 4.2.: **Modeling a genetic toggle switch.** (A) Schema of two mutually inhibiting genes  $x$  and  $y$ . (B) Mathematical model describing this interaction in terms of coupled ordinary differential equations. (C) Quasi-potential surface reconstructed from differential equations. Blue circles show (un)stable steady states, red line shows two possible trajectories from zero initial condition. (D) One-dimensional cut-out (along white line in (C)) depicting cell fate decision to either cell state expression either  $x$  or  $y$ .

When sufficient information on the transcriptional network is available, such a modeling approach can allow protocols for cellular reprogramming to be rationally designed, rather than obtained through trial-and-error, to efficiently convert a source cell type into a desired target cell type<sup>145</sup>. Using this strategy, Joseph Zhou and Lutz Brusch in our group, together with Sui Huang, constructed a hierarchical multi-attractor model based on the, incompletely known, pancreatic transcriptional network<sup>146</sup>. They showed this mathematical model was able to qualitatively reproduce the transcriptional dynamics during development and all pancreatic cell types. Subsequently, they used it to predict which genetic perturbations would result in the conversion into new  $\beta$ -cells, including the combination of factors that the Melton group had used<sup>146</sup>. Their finding showed that inhibition of certain transcription factors (Ptf1a), in combination with the overexpression of the Melton's cocktail, can improve reprogramming efficiency. Moreover, the sequence of overexpression was found to be important, optimal reprogramming was obtained with overexpression of MafA and Pdx1 first, and Ngn3 later<sup>146</sup>.

This proof-of-principle study showed that mathematical modeling of transcriptional regulation can provide important information to rationally design reprogramming experiments, despite the fact that little or no kinetic data is available on the regulation. Yet, this study was unable to account for several key aspects of pancreatic development. First, it was unable to reproduce the high acinar-to-islet cell ratio observed dur-

ing normal development where nascent islet cells are found scattered amidst acinar cells. On the one hand, this is due to the binary cell fate decisions resulting from the cross-inhibition between *Ptf1a* and *Ngn3* in the model. On the other hand, this is because such as single-cell model does not account for the known spatial interaction that governs the acinar-to-islet ratio, mediated by Notch signaling<sup>147,148</sup>.

#### 4.3 MATHEMATICAL MODELING OF INTERCELLULAR COMMUNICATION

So far, systems biological studies of cell fate regulation have mainly focused on regulation within genetic networks and have largely ignored the role of intercellular communication, in part due to the availability and dominance of omics data and the scarcity of spatiotemporal information. Yet, it is well-known that the prevalent mechanism of cell fate regulation in multicellular organisms is not cell-autonomous specification, but conditional specification that relies on tight coupling to signals from the microenvironment. Accounting for the multicellular context in which cellular decisions take place is therefore crucial to understand their regulation. By ignoring intercellular communication, such studies are unable to explain the spatiotemporal dynamics of cell fate control and the spatial regulation of ratios between cell types. For the design of reprogramming protocols, it causes one to overlook possibilities of cell type conversion strategies by manipulation of cell-cell communication rather than, more invasive and risky, genetic manipulations.

How communication between initially equivalent cells in a developing tissue results in the formation of complex patterns of distinct cell types has, of course, been extensively studied in the field of biological pattern formation. Theoretical work has shown that regional specification of cell fates can result from positional information<sup>69</sup>, typically consisting of a gradient of a diffusible signaling molecule, or from self-organized pattern formation based on the interaction between a short-range activator and long-range inhibitor species<sup>44,70</sup>. Fine-grained cellular patterns, on the other hand, typically emerge from contact-mediated communication between adjacent cells such as mediated by the Notch signaling pathway<sup>149–151</sup>.

The latter provides an instructive example. Consider a simple system of two cells inhibiting each others expression of a gene through a membrane-bound ligand-receptor mechanism (see figure 4.3A). In this mechanism, known as lateral inhibition, each cell has the possibility to be set in either of two expression states, depending on the state of the other cell. Interestingly, modeling this interaction in terms of ordinary differential equations for the two-cell system yields the identical system of equations as for the toggle switch (fig. 4.3B). Consequently, the steady states and potential surface are also identical. However, when solving this system of equations for a lattice of cells, we obtain a spatial salt-and-pepper pattern in which highly expressing cells are surrounded by low expressing cells (fig. 4.3C). If, however, the cell-cell communication is disrupted



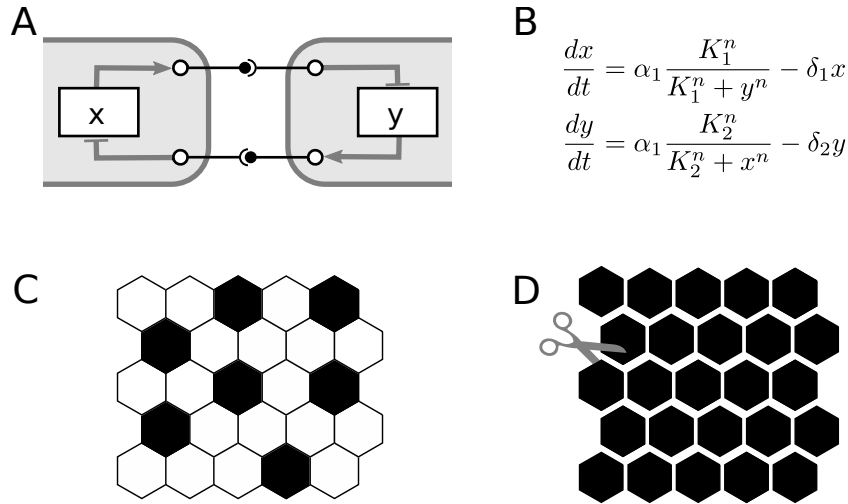


Figure 4.3.: **Modeling lateral inhibition.** (A) Schema of lateral inhibition between two cells with genes  $x$  and  $y$ , where  $x$  and  $y$  denote the same gene in two spatial locations. (B) Mathematical model describing the two-cell system (identical to figure 4.2B). (C) Salt-and-pepper pattern in a multicellular system showing highly expressing cells (black) inhibiting expression in neighboring cells (white). (D) After disruption of cell-cell communication, all cells show high expression in response to a lack of lateral inhibition.

through e.g. inhibition of signaling or tissue dissociation, all cells in the system will uniformly show high expression through the lack of inhibitory activity (fig. 4.3D).

This simple example shows that inhibition of Notch-mediated cell-cell signaling may be used to induce cell fate change and might, in some cases, offer alternatives to genetic manipulations. As a matter of fact, experiments have shown that inhibition of Notch signaling in the embryonic pancreatic epithelium results in an accelerated differentiation of nascent islet cells and a near absence of acinar cells<sup>147</sup>. Moreover, *in vitro* studies have shown that tissue dissociation induces spontaneous dedifferentiation of acinar cells<sup>152,153</sup> and inhibition of Notch signaling in these cells accelerates their conversion from acinar to  $\beta$ -cells<sup>154</sup>.

In general, there is a growing awareness of the importance of cell-cell communication mechanisms, also including cadherin and connexins, in stem cell biology and lineage conversion studies<sup>130,131,155–157</sup>. In addition to mathematical modeling of transcriptional networks, there is, therefore, a growing need for mathematical modeling approaches that take spatial cell-cell communication into account in order to understand induction and maintenance of pluripotency as well as to understand and improve methods of cellular reprogramming.

#### 4.4 CELL FATES, PATTERNS AND REPROGRAMMING IN THE PANCREAS

In the following chapters, cell fate regulation and pattern formation of exocrine (acinar) versus endocrine (islet) cell fates in the nascent pancreatic epithelium is studied. A novel mechanism of crosstalk between intercellular communication mechanisms is proposed in which lateral inhibition, mutual repression between adjacent cells, interacts with lateral stabilization, mutual induction between adjacent cells. Its consequences are explored with respect to pattern formation during pancreatic development (chapter 5) as well as to reprogramming strategies of converting pancreatic acinar cells into new islet cells (chapter 6).

In these studies, the piecemeal data from published reports on loss-of-function single-gene experiments is collected to construct a mathematical model based on tightly coupled ordinary differential equations (ODE) that qualitatively describes the expression of lineage-determining transcription factors through intracellular interactions and intercellular interactions between discrete cells (coupled ODE lattice). Bifurcation analysis of a minimal tissue, consisting of two or three cells, is used to reveal the different coexisting (multistable) patterning solutions by varying the relative strengths of mechanisms of intercellular communication. This analysis is complemented by simulations of larger systems to explore the effects of intercellular communication and noise on the control of spatial patterns, cell type ratios and efficiency of cell type conversion strategies. The results demonstrate that the scattered distribution of pro-endocrine cells in the developing pancreatic epithelium can arise through the proposed coupling of lateral inhibition and lateral stabilization, and that this also allows modulation of cell type ratios through the strength of lateral stabilization (chapter 5). Furthermore, it is shown that adult acinar and islet cell fates are multistable, opening the possibility of acinar-to-islet cell conversion without genetic manipulation, and that such conversion either involves dedifferentiation towards a progenitor-like multipotent state or proceeds by direct lineage switching, depending on the perturbation of intercellular communication (chapter 6).

#### 4.5 CONCLUSIONS

The theoretical and qualitative results presented in these studies are in line with experimental reports on the spatial and temporal expression of lineage-determining genes such as *Ngn3* and *Ptf1a* as well as with reported observations of “spontaneous” conversion of acinar cells upon enzymatic tissue dissociation. Additionally, and perhaps more importantly, these studies provide a formal framework to interpret and understand these experimental results based on a synergy between theories from pattern formation and cell fate regulation.

However, these studies have several shortcomings that are left for future investigation. For instance, the presented models do not take *cis*-interactions of ligands and

receptors into account, although the inhibiting effect of Delta ligands binding to Notch receptors on the same cell can have important effects on spatial patterning at the multicellular level<sup>158</sup>. Also, recent theoretical results show that Notch signaling can give rise to a third phenotype, with medium levels of ligand and receptors<sup>159,160</sup>. However, this requires the inclusion of a second ligand, Jagged, which has been omitted from the models presented here. Moreover, the identity of the molecules that involved the proposed lateral stabilization mechanism remain unclear. Despite indications that gap junctional communication, mediated by connexins, and cadherin-mediated signaling may be involved in this type of lateral signaling<sup>161,162</sup>, these suggestions are need to be validated experimentally. Finally, the modeling approach adopted in the following chapters ignores morphogenetic processes such as tissue growth and epithelial folding. However, since the specification of cell fates occurs simultaneously with these morphogenetic processes, changes in tissue shape may affect cell fate specification by some cells becoming more exposed to external signals. It has, for instance, been proposed that paracrine signals from the endothelium affect patterning of the pancreatic epithelium<sup>163,164</sup>. Investigating the effects of such morphogenetic changes and paracrine signals on spatial patterning and cell fate specification fits well within the modeling framework described in Part I, but is left for future work.

---

## ON THE ROLE OF LATERAL STABILIZATION DURING EARLY PATTERNING IN THE PANCREAS\*

---

### 5.1 INTRODUCTION

The pancreas is a complex organ consisting of two functionally distinct tissue compartments<sup>165</sup>. Exocrine acinar cells make up approximately 95-99% of cells in the pancreas and produce digestive enzymes that are released into the intestine. Cells organized in islets of Langerhans are endocrine and are vital in the regulation of glucose homeostasis throughout the body by releasing hormones such as insulin and glucagon into the blood. Both cell types are known to arise from a common pool of multipotent pancreatic precursors<sup>166–168</sup>. However, despite the identification of key transcription factors and intercellular signaling pathways, the mechanisms underlying the fate decision between these lineages remain unclear. In particular, the low endocrine-to-exocrine cell ratio and the scattered spatial distribution of early endocrine cells are poorly understood. Elucidation of these mechanisms may have important consequences for the development of therapeutic cell reprogramming and cell replacement therapies<sup>122</sup>.

Over the last decade, great progress has been made in revealing the transcriptional regulation of murine pancreatic development, and the endocrine compartment in particular<sup>169,170</sup>. Several transcription factors that are crucial for the cell fate decision between the exocrine and endocrine lineages have been identified (see<sup>169–171</sup> for reviews). Transgenic studies have identified neurogenin-3 (*Ng3*) as a pro-endocrine factor which is required to induce endocrine cell fates, since its overexpression results in massive conversion into endocrine cells at the expense of the exocrine compartment<sup>147,172</sup> and loss of *Ng3* causes depletion of endocrine cells<sup>173</sup>. Pancreas specific transcription factor 1 subunit alpha (*Ptf1a*) has been shown to be crucial for exocrine

---

\* This chapter is based on the publication: Walter de Back, Joseph Xu Zhou and Lutz Brusch, *Journal of the Royal Society Interface*, 10(79): 20120766, 2012. Walter de Back conceived the study together with Joseph Xu Zhou. Model construction, simulations and data analysis were performed by Walter de Back. Bifurcation analysis was done together with Lutz Brusch. Walter de Back wrote the paper.

specification, as mice deficient in this transcription factor form normal endocrine cells but lack an exocrine pancreas<sup>174–176</sup>.

Endocrine cells appear as individual cells or in small clusters scattered over the central pancreatic epithelium<sup>172,173,177,178</sup> with a low ratio of endocrine to exocrine cells, up to 1:50 in the developing pancreas, depending on the specific developmental stage. Investigation of intercellular signaling during pancreatic development (reviewed in<sup>170,179</sup>) suggests that local cell-cell communication, rather than long-range morphogen gradients, is important in lineage specification. In particular, it has been convincingly demonstrated that Notch signaling regulates the cell fate decision between exocrine and endocrine lineages. A series of transgenic studies shows that loss of genes in this pathway (*Dll1*, *Rbp-jk*, *Hes1*) causes precocious endocrine differentiation and loss of exocrine cells, similar to *Ngn3* misexpression<sup>147,148,154,179–182</sup>. This suggests that the exocrine-endocrine cell fate decision is controlled by contact-mediated lateral inhibition in which pro-endocrine cells inhibit endocrine specification of their neighboring cells, forcing them into an exocrine fate<sup>182</sup>.

Yet, the spatial distribution of endocrine cells within the early pancreatic epithelium is not well-explained by models of lateral inhibition<sup>179</sup>. Lateral inhibition alone causes the formation of fine-grained regular patterns, as observed in neural tissue and bristle patterns<sup>151,183,184</sup>, instead of the irregular scattered distribution of endocrine cells in the developing pancreas. Moreover, the predicted cell type ratio of 1:3 or 1:6 for lateral inhibition<sup>149</sup> does not agree with the observed endocrine to exocrine ratio in the pancreas. These inconsistencies suggest that additional mechanisms are involved in the pancreatic cell fate control.

An important clue can be found in studies of acinar-to-islet cell transdifferentiation *in vitro*, in which the endocrine-to-exocrine cell fate decision is recapitulated. In these experiments, adult acinar cells spontaneously de-differentiate to pancreatic progenitor-like cells upon dissociation of the tissue<sup>152,153,185,186</sup>. This suggests that maintenance of acinar cell fate crucially depends on a continuous signal provided by contacts between acinar cells<sup>187</sup>. Indeed, recent evidence shows that the dedifferentiation depends on the disruption of cadherin-mediated cell adhesion<sup>162</sup>. Similarly, inhibition of *Mist1* is reported to cause dedifferentiation due to disruption of gap junction intercellular communication<sup>188</sup>. Thus, cells of the exocrine pancreas appear to mutually stabilize their cell fate through physical contacts, in a mechanism that may be called lateral stabilization.

In this study, we propose that lateral stabilization provides positive feedback between pro-exocrine factors in adjacent progenitor cells and acts together with lateral inhibition in the regulation of lineage specification during early pancreas development. A simple mathematical model is constructed to capture the feedback mechanisms among pancreatic progenitor cells. Our analysis shows that the relative timing of the two feedback loops regulates the cell fate decision and tissue patterning in the central part of the developing pancreas. Specifically, our results show that the combi-

nation of lateral inhibition and lateral stabilization can explain the particular scattered spatial distribution of endocrine cells and provides means to regulate endocrine and exocrine cell type ratios in the pancreas.

## 5.2 METHODS

Interactions between transcription factor genes control cell fates by constraining the possible patterns of gene expression. Similarly, interactions between cells control patterning of a tissue through cell-cell signaling. The dynamics of gene-gene and cell-cell interactions can be modeled and analyzed in terms of differential equations. Whereas analysis of models of gene regulatory networks can reveal the existence of stable attractors that represent cellular phenotypes<sup>136,146,189</sup>, models of intercellular signaling, mediated by diffusive or membrane-bound ligands, can reveal pattern formation abilities in developing tissues<sup>70,149,150</sup>. In this study, the coupling between both modules, gene-gene interactions and cell-cell interactions, is analyzed mathematically to reveal the dynamics and attractors of gene expression and spatial patterning of endocrine cells in the pancreas.

The state of a cell is specified by two variables,  $X$  and  $Y$ , that represent the expression levels of fate-determining transcription factors.  $X$  represents a pro-endocrine transcription factor that is involved in lateral inhibition. Among the various pro-endocrine factors that have been described (*Nkx6.1*, *NeuroD*, *Ngn3*)<sup>169</sup>, only *Ngn3* is known to be actively involved in Notch signaling. *Ngn3* activates the expression of *Delta-like1* (*Dll1*)<sup>147</sup> and is suppressed by *Hes1* upon *Notch* receptor activation<sup>148</sup> (figure 5.1A). Therefore, we interpret  $X$  as the expression level of the transcription factor *Ngn3*.  $Y$  represents a factor that is expressed in both progenitor and exocrine cells, but inhibited in cells that commit to the endocrine lineage. Therefore,  $Y$  is interpreted as the transcription factor *Ptf1a* because this is the only factor with that specific expression profile and known to be necessary and sufficient to induce the exocrine cell fate<sup>174,176</sup>.

We assume a weak external activation,  $\theta$ , for both  $X$  and  $Y$ . This is based on evidence that *Ngn3* is activated by *Hnf6*<sup>190</sup> and *Ptf1a* is activated by *Hnf1 $\beta$* <sup>191</sup> which is itself regulated by *Hnf6*<sup>192</sup>. For the sake of simplicity, the external activation is assumed to be constant during the developmental stage.

Cells in our model interact with adjacent cells through two cell-cell signaling mechanisms: lateral inhibition and lateral stabilization. The factor  $X$  in each cell mediates lateral inhibition of surrounding cells: the rate of production of  $X$  is down-regulated by expression of this factor in neighboring cells. This mechanistically captures the well-established pathway that expression of *Ngn3* upregulates the Notch ligand *Dll1* which, when bound to Notch receptors on adjacent cells, activates the expression of *Hes1* which represses *Ngn3* in these adjacent cells<sup>147,148,181,193</sup>.

Factor  $Y$  is involved in lateral stabilization which provides a positive feedback loop between  $Y$ -expressing neighboring cells. The rate of production of  $Y$  is up-regulated by

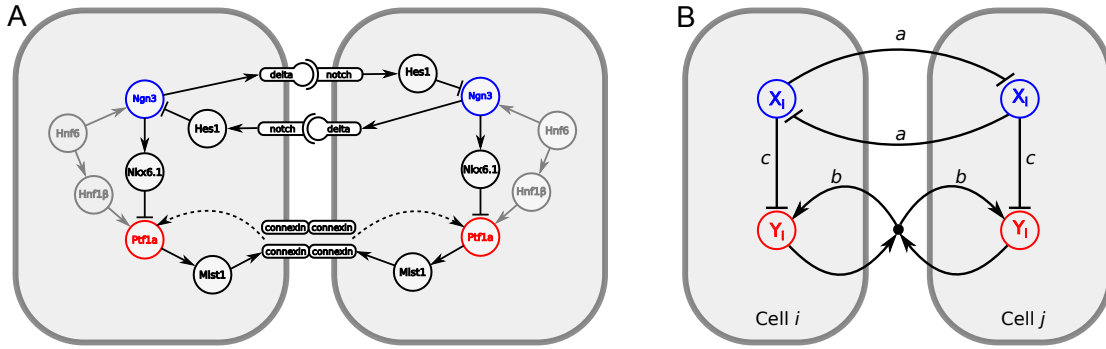


Figure 5.1.: **Interactions between transcription factors and signaling pathways.** (A) Known regulatory interactions involved in the exocrine-endocrine cell fate decision in the pancreas, including contact-mediated signaling (see main text for details and references). Receptor-ligand binding in the Notch signaling pathway induces lateral inhibition. Formation of gap junctions represents one possible pathway for lateral stabilization (dashed arrow). (B) Interactions in the two-variable model in which cells  $i$  and  $j$  are coupled by lateral inhibition and lateral stabilization (see main text for details). Parameters  $a$ ,  $b$ , and  $c$  represent the interaction strengths.

simultaneous expression of  $Y$  in neighboring cells. Although the molecular details of a lateral stabilization pathway are unclear, such conditional activation is in principle consistent with both cadherin/ $\beta$ -catenin signaling<sup>162</sup> and gap junctional communication via expression of *Mist1* and *connexin*, downstream targets of *Ptf1a*<sup>188</sup>. In both cases, the expression of gene products in the form of homotypic transmembrane molecules is required in all participating cells to allow intercellular signaling. Mathematically, this is captured by a multiplication such that  $Y^-$  (non-expressing) cells do not contribute nor benefit from stabilization.

Despite many observations that endocrine markers (e.g. *Isl1*) and *Ptf1a* are mutually exclusive<sup>194</sup>, the underlying regulatory mechanisms remain unclear. One proposal states that *Nkx6.1*, a pro-endocrine factor downstream of *Ngn3*<sup>195,196</sup>, antagonizes the expression of *Ptf1a*<sup>197</sup>. Independent of the precise molecular pathway, we can assume that pro-endocrine factors suppress the expression of *Ptf1a* leading to the restriction of the latter factor to the exocrine compartment. Accordingly,  $X$  inhibits  $Y$  in a cell-autonomous fashion in our model.

At a later stage during pancreas development, *Hnf6* and *Ngn3* become downregulated and are not expressed in the adult pancreas. Because our primary interest lies in the lineage specification prior to this stage, we can neglect this downregulation. This considerably simplifies the model formulation and analysis, but implies that the common endocrine marker *Ngn3* cannot be used to indicate commitment to the endocrine lineage. Instead, the absence of *Ptf1a* (i.e.  $Y$ ) expression will be used to mark endocrine cells, as supported by single-cell transcript analysis<sup>194</sup>.

In this fashion, the details of the molecular pathways (fig. 5.1A) have been reduced to a core regulatory interaction network (fig. 5.1B) which has been formalized in terms of stochastic ordinary differential equations using Hill kinetics as follows (see table 5.1):

$$\begin{aligned}\frac{dX}{dt} &= \frac{\theta^n}{\theta^n + a\bar{X}^n} - X + \xi_X(t) \\ \frac{dY}{dt} &= \frac{\theta^n + b(Y\bar{Y})^n}{\theta^n + b(Y\bar{Y})^n + cX^n} - Y + \xi_Y(t).\end{aligned}\tag{5.1}$$

Both transcription factors  $X$  and  $Y$  are activated by  $\theta$ , which is chosen small relative to the parameters of cell-cell interaction. The parameter  $a$  represents the strength of lateral inhibition by neighboring cells,  $b$  denotes the strength of lateral stabilization and  $c$  models the strength of cell-autonomous inhibition of  $Y$  by  $X$ . To focus on the impact of lateral stabilization, the model is scaled such that external activation  $\theta$  is weak compared to the interaction terms and the parameters  $a$  and  $c$  have been set to unity. In this way, the strength of lateral stabilization  $b$  can be treated as a control parameter. Both factors are subjected to non-regulated first order degradation, such that expression levels are between 0 (not expressed) and 1 (fully expressed). The terms  $\bar{X}$  and  $\bar{Y}$  denote the average expression of  $X$  and  $Y$  in neighboring cells. The additive stochastic terms  $\xi(t)$  are random variables with a Gaussian white noise distribution  $N(0, \eta)$  with mean 0 and amplitude  $\eta$ . The Hill coefficient  $n$  is chosen to achieve non-linear step-like behavior ( $n = 4$ ). Production of  $X$  is inhibited by the expression of  $X$  in neighboring cells,  $a\bar{X}^n$ , independent of its own activation. In contrast, cell-autonomous activity of  $Y$  is required for an increase in production of  $Y$  by the lateral stabilization term,  $b(Y\bar{Y})^n$ . Thus, this cell-cell interaction acts to stabilize a pre-existing expression.

Pattern formation abilities are determined by the topology of the network, rather than the precise kinetic parameters<sup>146,198</sup>. Thus, for the purpose of this study, the molecular details of the regulatory pathways can be lumped into activatory or inhibitory interactions without altering the qualitative behavior of the system.

This qualitative treatment allows us to work with a small enough number of variables and parameters to gain insight into the spatiotemporal dynamics by mathematical analysis. Although the various patterning solutions reported below are independent of our particular choice of parameters (table 5.1), they do depend on the Hill coefficient  $n > 2$  to induce the required bistability in  $X$ .

Concerning initial conditions, it is known that expression of *Ngn3* and *Ptf1a* are initiated around the same time in the developing pancreas. Both factors are, directly or indirectly, induced by the same upstream transcription factor *Hnf6* and are both first detected around E9 in mice<sup>175</sup>. Therefore, we consider the simultaneous initiation of  $X$  and  $Y$  ( $(X, Y) = (0, 0)$  at  $t = 0$ ) by external activation at rate  $\theta$ . We assume a symmetric activation of both factors ( $\theta_X = \theta_Y$ ) for reasons of clarity, but unequal acti-



Table 5.1.: Variables, parameters and observables of the mathematical model (equation 5.1).

	Symbol	Description	Value
<b>Variables</b>	$X$	Expression of pro-endocrine transcription factor <i>Ngng3</i>	0 (initial)
	$Y$	Expression of transcription factor <i>Ptf1a</i>	0 (initial)
	$\bar{X}$	Average <i>Ngng3</i> expression in neighboring cells	0 (initial)
	$\bar{Y}$	Average <i>Ptf1a</i> expression in neighboring cells	0 (initial)
<b>Parameters</b>	$a$	Strength of lateral inhibition	1
	$b$	Strength of lateral stabilization	20
	$c$	Strength of intracellular inhibition	1
	$\theta$	External activation	0.1
	$n$	Hill coefficient	4
	$\eta_x$	Amplitude of Gaussian white noise on $X$	$10^{-4}$
	$\eta_y$	Amplitude of Gaussian white noise on $Y$	0
<b>Observables</b>	$\tau_x$	Time until symmetry break in $X$	
	$\tau_y$	Time until super-induction of $Y$	
	$\epsilon$	Ratio of endocrine cells	

vation ( $\theta_X \neq \theta_Y$ ) results in the same qualitative behavior. The numerical simulations presented below use this zero initial condition.

Morphogenetic events such as proliferation, cell motility, branching are excluded from the model. Lineage tracing experiments have shown that the majority of cells are committed to the exocrine or endocrine compartments around E10, before the onset of major morphogenetic events<sup>166,199</sup>. In the relevant developmental stage, the pancreatic epithelium can thus roughly be approximated by a fixed two-dimensional lattice of densely packed hexagonal cells.

Analysis and numerical simulation were performed using GRIND (phase plane analysis)<sup>200</sup>, XPPAUT (bifurcation analysis)<sup>201</sup> and our modeling environment Morpheus (lattice simulations)<sup>55</sup> (see chapter 3). Numerical integration was performed using the Runge-Kutta (RK4) method with time step size  $dt = 0.02$ . The deterministic model of the cell couplet is available in SBML format is available in the BioModels database\*. The full description of the simulation model is listed in MorpheusML in figure 3.7 and is available in MorpheusML format in the Supplementary Online Material under <http://walter.deback.net/thesis>.

To explore the effects of different source of noise (see section 5.3.5 and fig. 5.5), simulations were conducted for different stochastic ODEs:  $\frac{dX}{dt} = \frac{\theta^n}{\theta^n + \xi_a(t) a \bar{X}} - X + \xi_x(t)$  where the stochastic terms  $\xi_x(t)$  (gene expression) and  $\xi_a(t)$  (signaling) are random variables with a Gaussian white noise distribution  $\xi(t) = N(0, \eta)$  with mean 0 and amplitude  $\eta_x$  and  $\eta_a$  resp. In the prepatter model, deterministic equations are used with non-homogeneous initial condition  $X(t = 0) = \xi_0$  where  $\xi_0$  is an exponential distribution  $\xi_0 = E(\eta_0)$  with amplitude  $\eta_0$ . Initial conditions with irregular cell shapes and contacts were set up as using the cellular Potts model (CPM)<sup>79</sup> employing a modified Metropolis algorithm to model motility of cell boundaries (see equation 9.3) where  $T$  controls the amplitude of fluctuations under the Hamiltonian  $H = I + \sum_{\sigma > 0} (a_\sigma - A_\sigma)^2$  which constraints cells to retain an area  $A$  while reducing the number of cell-cell interfaces  $I$ . CPM-specific parameters were chosen as  $A = 60$ ,  $Y = 0.01$ . The observable time  $\tau_X$  represents the point at which  $X$  expression between neighboring cells have diverged above a threshold value ( $\|X(\tau_X) - \bar{X}(\tau_X)\| \geq 0.25$ ), and  $\tau_Y$  represent the point when  $Y$  expression exceeds a threshold value ( $Y(\tau_Y) > 0.75$  or  $\bar{Y}(\tau_Y) > 0.75$ ).

### 5.3 RESULTS AND DISCUSSION

To investigate the pattern formation properties, the model was computed on a lattice under varying conditions of cell-cell signaling. The results in figure 5.2 demonstrate that various spatial patterns can arise under different combinations of cell-cell interaction parameters.

\* <http://www.ebi.ac.uk/biomodels>, BioModels ID: MODEL1211010000

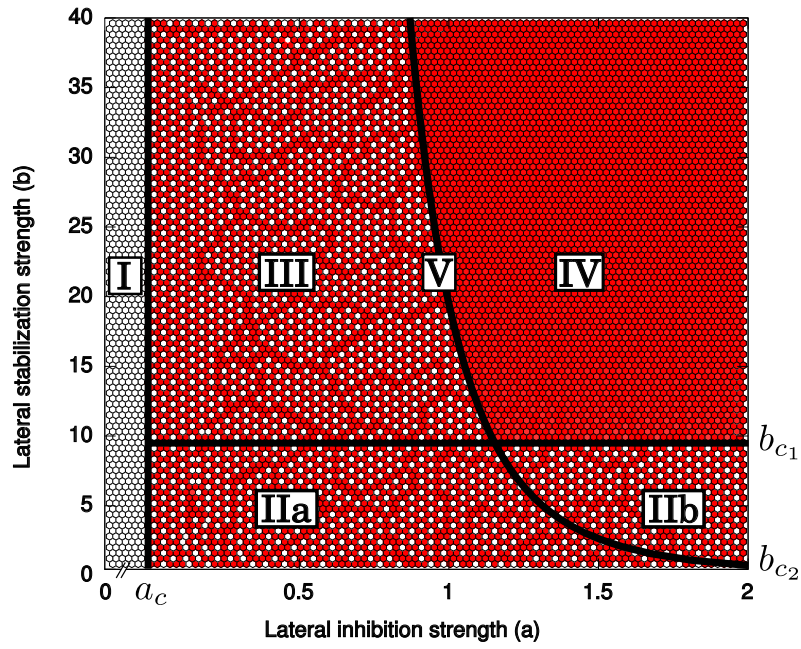


Figure 5.2.: **Classification of different spatial pattern domains under varying strengths of lateral inhibition  $a$  and lateral stabilization  $b$ .** Color (dark/white) indicate (high-/low) expression of  $Y$ . The Roman numerals (I...V) denote different patterning domains. Solid lines indicate phase transitions with critical values  $a_c$ ,  $b_{c1}$  and  $b_{c2}$  that are found by means of bifurcation analysis. Parameters as in table 5.1,  $a$  and  $b$  as indicated on axes.

### 5.3.1 Precocious endocrine specification

Domain I in figure 5.2 shows the absence of exocrine cells under conditions of weak or no lateral inhibition,  $a < a_c$  where  $a_c = \theta^n$ , independent of the strength of lateral stabilization  $b$ . This is in line with seminal loss-of-function studies showing that deficiency in Notch signaling pathway members causes precocious endocrine differentiation and absence of *Ptf1a*-expression<sup>147,148</sup>. In the context of the model, the reasons behind this phenotype are straightforward. In absence of lateral inhibition, expression of  $X$  is uninhibited and exhibits simple saturated growth. Due to intracellular antagonization of  $X$  on  $Y$ , the upregulation of the pro-exocrine factor is always slower, independent of the strength of lateral stabilization. Therefore, the fast uninhibited expression of  $X$  results in (1) fast commitment to the endocrine lineages and (2) absence of exocrine cells due to the inhibition of  $Y$ .

### 5.3.2 Checkerboard patterning

Over a wide range of parameters, the system organizes itself into a regular well-spaced, checkerboard-like pattern (figure 5.2). This spatial distribution is expected for models of lateral inhibition<sup>149,151</sup> and known to arise as a result of Notch signaling during neurogenesis in *Drosophila* as well as in vertebrates<sup>202,203</sup>. It is instructive to consider the dynamics of  $X$  in a system of only two cells  $i$  and  $j$  for which the phase plane is depicted in figure 5.3A,B. This shows the coexistence of an unstable steady state where  $X_i = X_j$  and two stable steady states where  $X_i \neq X_j$ . From the zero initial condition, the dynamics of expression displays a rapid initial evolution towards the unstable equilibrium (arrowhead in figure 5.3C). This represents the expression of  $X$  in the undecided progenitor state, the level of which inversely depends on the strength of lateral inhibition  $a$ . Subsequently, perturbations around this symmetric unstable state self-amplify and result in the divergence into opposite states of expression. The timing of the break of symmetry,  $\tau_x$  (arrow in figure 5.3C), depends on the amplitude of noise  $\eta$ , because, on average, critical perturbations arise earlier under increased levels of stochasticity. In a two-dimensional array of cells, lateral inhibition causes the establishment of alternating expression states<sup>149</sup> following the propagation of the breaking of symmetry. Defects in regularity can occur at the boundary of those domains due off-register initial symmetry breaks in distant cells.

In all domains where checkerboard patterning occurs (IIa, IIb and III), the expression of  $Y$  is “enslaved” through the antagonization by  $X$ . In domain II(a and b), the strengthening effect of lateral stabilization is too weak to counteract intracellular inhibition by  $X$ . The reasons for the same type of patterning in domain III are more subtle and discussed below.

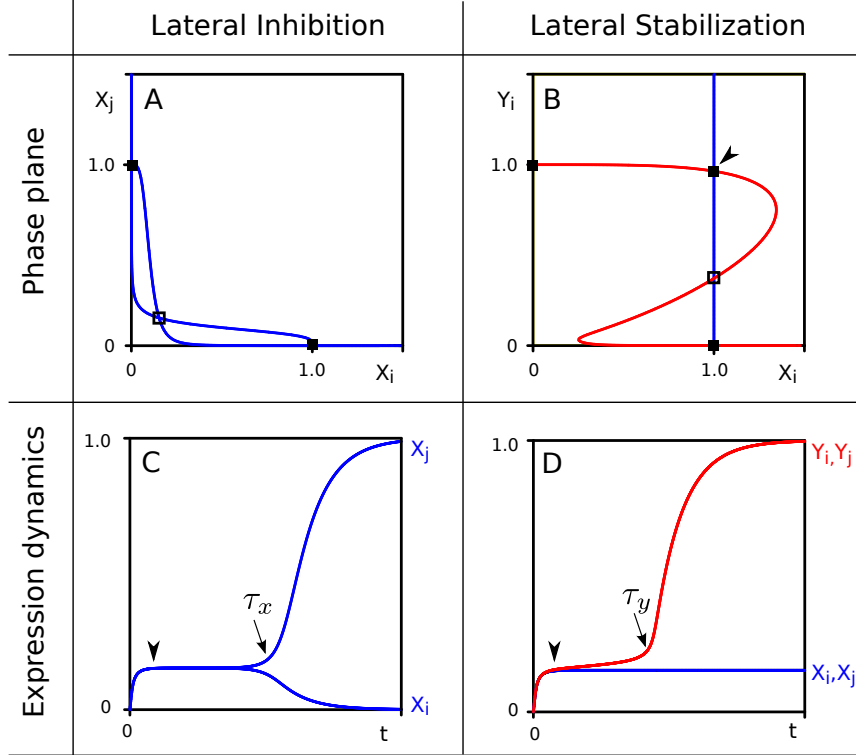


Figure 5.3.: **Analysis of cell-cell interaction mechanisms in a system of two cells  $i$  and  $j$ .** Phase plane analysis (A,B) and time plots of gene expression dynamics (C,D). (A) Nullclines of  $X$  (where  $dX/dt = 0$ ) under lateral inhibition. One symmetric unstable steady state (open box) and two asymmetric stable steady states (filled boxes) coexist. (B) Nullclines of  $X_i$  (solid, blue) and  $Y_i$  (dashed, red) under lateral stabilization where  $Y_j = 1$ . One unstable steady state (open box) and three stable steady states (filled boxes) exist. Arrowhead indicates 'new' steady state dependent on lateral stabilization. (C) Symmetry breaking in  $X$  expression. After induction, a transient of intermediate expression at the symmetric unstable steady state (arrowhead) is followed by symmetry-breaking (arrow, at  $t = \tau_x$ ) into one  $X^+$  and one  $X^-$  cell. (D) Biphasic growth of  $Y$  in the deterministic system ( $\eta = 0$  to exclude symmetry breaking of  $X$ ). After induction, expression of  $Y$  remains at a plateau level (arrowhead) followed by super-induction (arrow, at  $t = \tau_y$ ) through lateral stabilization. Parameter values as in table 5.1.

### 5.3.3 Biphasic growth through lateral stabilization

A homogeneous pattern of cells expressing the pro-exocrine Y factor is observed in domain IV (figure 5.2). That is, when both intercellular signaling mechanisms are strong, an absence of endocrine cells is observed. Under these conditions, the positive feedback between Y expressing cells is strong enough to escape antagonization by X. Indeed, the phase portrait given in (figure 5.3B) shows the existence of stable attractors where Y is highly expressed, independent of X.

Interestingly, the dynamics of Y towards this expression state show biphasic growth in which phases of growth are separated by a transient plateau (arrowhead in figure 5.3D). Upon initial activation, there is initial growth up to a level where it is approximately balanced by suppression by X, which is itself expressed at intermediate levels. At this stage, both cells have moderate expression of Y and mutually promote each other's expression of Y through positive feedback. If this feedback is strong enough, it accelerates expression and initiates the second growth phase. The timing of the superinduction at  $t = \tau_Y$  (figure 5.3D) and whether it occurs at all, depends on the strength of the stabilizing coupling between cells (b).

Intriguingly, the prediction of biphasic growth is in line with the expression pattern of *Ptf1a* during pancreatic development. While this factor is maintained at low expression levels at the progenitor stage, *Ptf1a* expression is superinduced in cells initiating acinar cell differentiation<sup>204</sup>.

The fact that strong lateral inhibition is necessary for ubiquitous exocrine specification seems counterintuitive, since lateral inhibition is responsible for the establishment of a heterogeneous checkerboard patterning. However, as mentioned, the strength of lateral inhibition  $\alpha$  determines the level of intermediate expression in the uncommitted progenitor state and thereby the inhibition of Y. Accordingly, weak lateral inhibition leads to high expression levels of X in the progenitor state and therefore inactivates the lateral stabilization mechanism of exocrine specification, resulting in a reduction of exocrine cells (domain III).

### 5.3.4 Multi-stability of patterns

To investigate how the stability of different patterning solutions changes under the influence of model parameters, we performed a bifurcation analysis on a reduced system of two cells. Figure 5.4A shows a bifurcation diagram where the strength of lateral stabilization  $b$  is varied, while keeping lateral inhibition strength constant,  $\alpha = 1$ . Comparing with figure 5.2, we observe that the system passes through three qualitatively different domains (Ia, III, IV) and two spatial patterning solutions: a checkerboard and a 'homogeneous' pattern. Whereas in the checkerboard pattern, the endocrine and exocrine cell fates are found in an alternating pattern, only exocrine  $Y^+$  cells exist in the 'homogeneous' pattern, despite small differences in Y expression. Two critical values

$b_{c_1}$  and  $b_{c_2}$  are identified that determine the transition between domains. Below  $b_{c_1}$ , the only stable pattern is the checkerboard solution ( $Y_i + Y_j = 1$ ). Indeed, the checkerboard solution is stable for all  $b$ . Between the critical values  $b_{c_1} < b < b_{c_2}$ , a stable homogeneous solution of high  $Y$  expression coexists, but this cannot be reached from zero initial conditions. This is because the expression at the plateau is stable for  $Y$  but not for  $X$  (in red in figure 5.4). Thus, expression of  $Y$  rests at an intermediate level and ‘waits’ until the symmetry in  $X$  expression is broken after which the dynamics of  $Y$  become enslaved by  $X$  ( $\tau_x < \tau_y$ , figure 5.6A). Conversely, for  $b > b_{c_2}$ , the stable state at plateau-level expression does not exist. Instead of resting at intermediate levels,  $Y$  expression continually grows as a result of lateral stabilization. This causes super-induction of  $Y$  before the symmetry of  $X$  is broken and thus, a homogeneous exocrine pattern emerges ( $\tau_x > \tau_y$ , figure 5.6B).

These results demonstrate that (1) the model exhibits multi-stability of patterning solution for a wide range of parameters ( $b > b_{c_1}$ ) and (2) under initial conditions relevant for development, the deterministic model switches from a heterogeneous (checkerboard) to a homogeneously exocrine pattern when lateral stabilization strength exceeds the critical value  $b_{c_2}$ . This switch in spatial patterning is driven by the relative timing between lateral inhibition and lateral stabilization mechanisms.

### 5.3.5 Noise controls timing of cell fate decision

Next, we studied the effect of noise on this transition. Because the timing of the cell fate decision due to lateral inhibition,  $\tau_x$ , depends on the amplitude of noise  $\eta$ , an increase in stochastic fluctuations will trigger a faster emergence of  $X^+$  and  $X^-$  cells. In contrast, the mechanism of lateral stabilization is, by itself, insensitive to stochasticity, and in fact acts to homogenize noisy expression<sup>150</sup>. Autonomously, the time  $\tau_y$  required for the lateral stabilization feedback loop to exceed a critical threshold value depends only on its strength  $b$ . Yet, when coupled to the lateral inhibition module upstream, stochasticity will also affect lateral stabilization. When, for a particular noise amplitude, the time to break symmetry in  $X$  decreases below the time required for lateral stabilization, i.e.  $\tau_x < \tau_y$ , the dynamics of  $Y$  become enslaved to  $X$ . Consequently, less exocrine and more endocrine cells arise, thus increasing the endocrine cell ratio (see figure 5.5).

Under noisy conditions, a checkerboard pattern may arise, even in the case  $b > b_{c_2}$ , if a fluctuation happens to lead to a fast symmetry break of  $X$  which enslaves the dynamics of  $Y$ . Conversely, a homogeneous  $Y^+$  solution may emerge for  $b_{c_1} < b < b_{c_2}$  if, by chance, the divergence of  $X$  is delayed which allows the super-induction of  $Y$  by stabilization. Thus, the discontinuous phase transition at  $b_{c_2}$  becomes continuous in the presence of noise due to its modulation of the timing in expression dynamics.

Many sources of cellular noise exist that may affect cell fate decisions. In addition to noise in gene expression, e.g. due to transcriptional bursts, there are stochastic factors

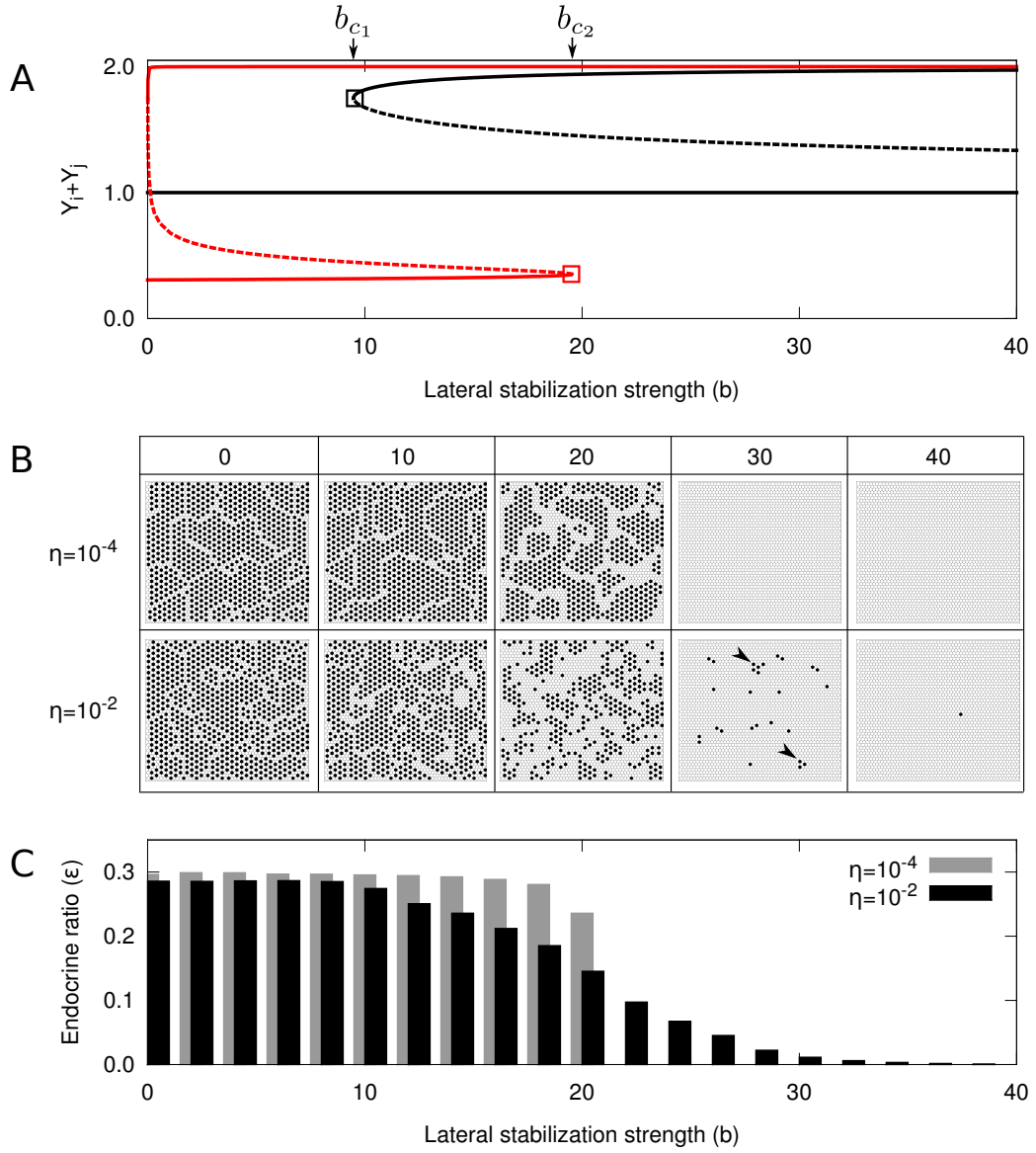


Figure 5.4.: **Scattering and cell type ratio control by lateral stabilization.** (Caption on next page.)



Figure 5.4.: **Scattering and cell type ratio control by lateral stabilization.** (A) Bifurcation diagram showing the steady states of  $Y$  expression of a two-cell system. Two saddle-node bifurcations are found at  $b = b_{c1}, b_{c2}$  that explain the phase transitions in figure 5.2. The four-variable system is projected onto a one-dimensional scale using the sum of  $Y$  expression in the couplet ( $Y_i + Y_j$ ). Dashed lines indicate unstable states, solid lines are stable. Black lines represent globally stable states, thick (red) lines represents states in which  $Y$  is stable but  $X$  is unstable. (B) Spatial patterns of endocrine (black) and exocrine (white) cells in lattice simulation at different strength of lateral stabilization  $b$  and under two conditions of noise  $\eta$ . Under higher noise levels, scattering is observed at the phase transition (see domain V in figure 5.2). Arrowheads indicate small clusters of checkerboard pattern. (C) Cell type ratio of endocrine cells  $\epsilon$  as a function of lateral stabilization strength  $b$ . Grey/black indicate noise levels. The simulation model is available in the Supplementary Online Material under <http://walter.deback.net/thesis>.

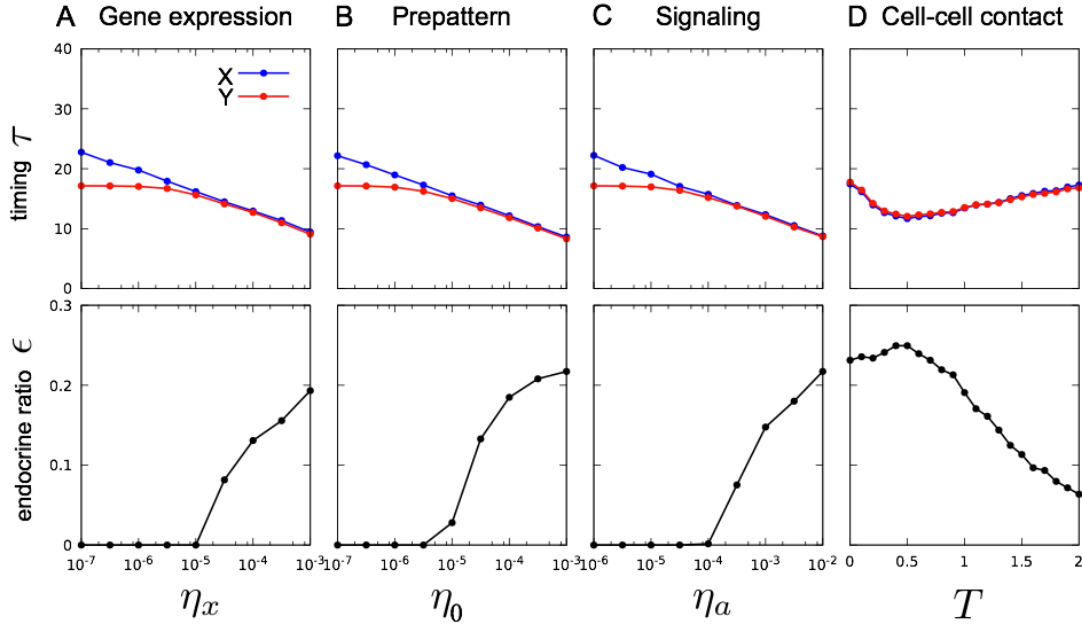


Figure 5.5.: **Effect of different sources of noise on timing and endocrine cell ratio.** Higher noise amplitudes decrease time to decision and increase endocrine cell ratio for different sources of noise (A, B, C). Small irregularities in local cell-cell contacts (D) further increases the endocrine cell ratio by allowing faster breaking of symmetry between neighboring cells. However, when cell-cell contact irregularities are large enough to include contact between non-local neighbors, the endocrine ratio decreases (D), in line with observations of pattern refinement via filopodia in *Drosophila*<sup>205</sup>. Simulations are performed as described in section 5.2. Parameters are as in table 1, except that  $b = 21$  such that the deterministic equations are in domain IV. Points represent averages over 10 independent simulations.

influencing signal transduction as well as fluctuations of local cell-cell contacts. To investigate the effects of different sources of noise, we systematically explored different noise terms in the ODE model. The results show that preexisting cell-to-cell variability or noise on signaling have qualitatively identical effects as gene expression noise (cf. fig. 5.5A-C). Interestingly, however, irregularities in local cell-cell contacts can have the opposite effect if contacts between non-neighboring cells occur. In this case, slower symmetry breaking and a reduction of the endocrine cell ratio is observed (fig. 5.5D), in line with observations of pattern refinement via filopodia in *Drosophila*<sup>205</sup>. Thus, both molecular and cell-level sources of variability are important in control of the timing of cell fate decisions and the regulation of cell type ratios. Note that the same does not hold for the phase transition at  $b_{c1}$ , which is independent of  $X$  and therefore insensitive to noise.

### 5.3.6 Scattering of endocrine cells under noise

To investigate the effect of stochasticity on patterning, simulations were performed on a lattice of cells. Figure 5.4B shows that under low noise levels ( $\eta = 10^{-4}$ ), as expected, a checkerboard pattern of pro-endocrine cells is observed for  $b \leq b_{c2}$ , whereas a homogeneous exocrine pattern is established for  $b > b_{c2}$ . In contrast, under higher noise amplitude ( $\eta = 10^{-2}$ ), a continuous transition appears for  $b > b_{c2}$  that is characterized by the concomitant appearance of the two multi-stable patterns in the same tissue. In other words, noise expands the region in parameter space where checkerboard and homogeneous exocrine patterning may occur side by side (domain V in figure 5.2).

Above the critical value  $b_{c2}$ , we observe the emergence of a scattered distribution of endocrine  $Y^-$  cells amidst a majority of exocrine  $Y^+$  cells. The particular scattering pattern predicted by our model is characterized by the presence of isolated or small clusters of endocrine cells, consisting of local patches of a checkerboard pattern (arrow-heads in figure 5.4B,  $b = 30$ ). Although no detailed study on the spatial distribution of endocrine cells during pancreas development has been conducted to our knowledge, the observed pattern is in agreement with the numerous reports on a scattered distribution in which endocrine cells were found as individual cells or small clusters in the pancreatic epithelium<sup>147,148,172,173,177</sup>.

In the model, the formation of these clusters involves a competition between two wave fronts. Once an early  $X^+$  cell arises by chance, it initiates the propagation of the checkerboard pattern by inhibition of  $X$  in neighboring cells. While inhibiting  $X$  in adjacent cells, these become  $Y^+$  and in turn initiate the propagation of a second wave by stabilization of  $Y$  in their neighbors. When the wave of lateral stabilization propagates faster than the wave of lateral inhibition, the checkerboard pattern halts at a particular size. Thus, the model predicts the propagation of an inductive signal through homeogenetic induction, in which a differentiated cell causes the differentiation of nearby progenitor cells. To date, no evidence of homeogenetic signals exists for the

developing pancreas. However, novel dynamic imaging studies may provide a more detailed insight into the spatiotemporal progression of signals and the precise spatial distribution of endocrine cells.

### 5.3.7 Lateral stabilization modulates cell type ratios

An interesting consequence of these dynamics is that it allows the regulation of cell type ratios in the pancreas, even in absence of proliferation. Due to noise in expression, endocrine-to-exocrine cell type ratios can be modulated between the ratios expected for checkerboard patterning by lateral inhibition (1:3) and the homogeneous exocrine solution (0:1) (see figure 5.4C). For a particular set of parameters, the locations where clusters of endocrine cells appear are chosen randomly but the ratio of cells that commit to the endocrine and exocrine lineage is determined by the parameters and can thus be predicted. From this model, it follows that down-regulation of the strength of lateral stabilization will dramatically increase the endocrine to exocrine cell type ratio.

Although proliferation is clearly important in the regulation of pancreatic cell type ratios during development, several experimental studies are in agreement with this finding. For instance, it was found that a conditional knock-out of  $\beta$ -catenin downregulates *Ptf1a* which results in a striking paucity of exocrine acinar cells, while preserving the endocrine compartment<sup>206,207</sup>. Conversely, overexpression of  $\beta$ -catenin around E12.5 leads to an increased exocrine cell mass, with only minimal changes to the endocrine cell count<sup>208</sup>. Although  $\beta$ -catenin is a key player in Wnt signaling as well as in cadherin-mediated cell-cell contact and it remains unclear which role is decisive in the context of cell type ratio control<sup>207</sup>, it is tempting to assume its participation in lateral stabilization of *Ptf1a* through cell-cell contacts takes part in the regulation of cell type ratios in the pancreas.

### 5.3.8 Related mechanisms

It should be noted that the proposed coupling between lateral inhibition and lateral stabilization is not the only mechanism that can produce scattering patterns. In particular, models using an activator-inhibitor mechanism that include a long-range inhibitor produced by cells entering one lineage can generate scattered distributions, even without Notch-mediated lateral inhibition<sup>74</sup>. However, although the mesenchyme surrounding the early pancreas emits diffusive inhibitory signals<sup>209,210</sup>, no evidence exist to suggest that such signals are produced within the primitive pancreatic epithelium itself. In contrast, there is accumulating evidence that short-range contact-mediated signals are crucial for lineage regulation in the pancreas<sup>154,162,188</sup>, and the role of Notch signaling is already well-established<sup>147,148,211</sup>.

The mechanism of lateral stabilization proposed in this study is closely related to other inductive mechanisms. Similar to lateral stabilization, the community effect is an

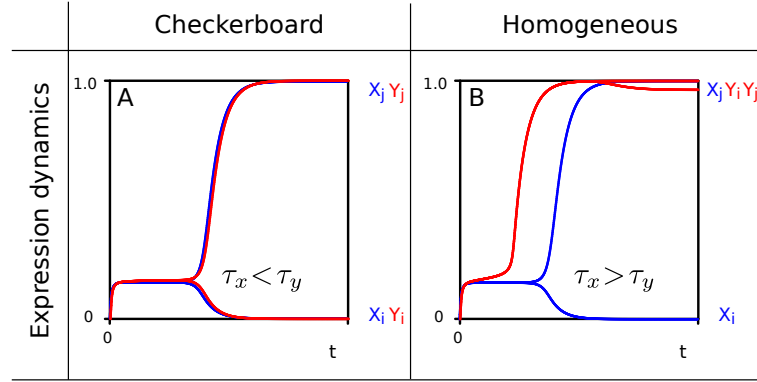


Figure 5.6.: **Relative timing of cell fate decisions depends on lateral stabilization ( $b$ ) and controls patterning.** (A) For  $b \leq b_{c_2}$ , symmetry-breaking of  $X$  precedes super-induction of  $Y$  ( $\tau_x < \tau_y$ ), resulting in checkerboard patterning ( $b = 15$ ). (B) In contrast, for  $b > b_{c_2}$ , super-induction of  $Y$  precedes symmetry breaking of  $X$  ( $\tau_y < \tau_x$ ) and results in ubiquitous exocrine cell fates ( $b = 35$ ).

inductive mechanism that results in the homogenization of cell fates<sup>150,212</sup>. However, it is typically associated with diffusive signaling among large groups of cells<sup>213</sup>, whereas the contact-mediated mechanism of lateral stabilization only requires small groups of cells. Moreover, the community effect operates by receptor-ligand binding<sup>214</sup>, whereas lateral stabilization assumes homodimerization of membrane-bound proteins. This implies that for lateral stabilization to take place, all participating cells must already be alike in their expression of this particular protein. In this sense, lateral stabilization also bears similarity with the homeogenetic induction mechanism introduced by Mangold and Spemann<sup>215</sup> in which differentiated cells induce surrounding undifferentiated cells to commit to the same lineage. Although the presence of differentiated cells is not required for lateral stabilization because interaction between progenitor cells themselves can lead to uniform induction (if  $b > b_{c_2}$ ), their presence would set out a wave of induction which eventually homogenizes cell fates in the tissue (even if  $b_{c_1} < b \leq b_{c_2}$ ). In fact, this predicted wave propagation effect can be used to experimentally test the lateral stabilization mechanism using grafting of adult exocrine ( $Ptf1a^+$ ) cells within a population of pancreatic progenitor cells.

#### 5.4 CONCLUSION

Despite indisputable evidence that Notch signaling is important in the regulation of the cell fate decision between the endocrine and exocrine lineages<sup>147,148</sup>, it is still controversial whether it acts through lateral inhibition or an alternative mechanism such as suppressive maintenance<sup>211,216–218</sup>. Yet, none of the theories proposed to date is able to explain the observed scattered distribution of endocrine cells in the pancreatic ep-

ithelium<sup>179</sup>. Interestingly, recent evidence from transdifferentiation studies suggests the existence of an additional mechanism that stabilizes pancreatic exocrine cells fate by a positive feedback loop mediated by cell-cell contacts. In this study, we have shown that the coupling between lateral inhibition and lateral stabilization may control early lineage specification and patterning in the pancreas. We have constructed and analyzed a minimal mathematical model combining these contact-mediated signalling mechanisms and showed that its behavior agrees with recent transgenic loss- and gain-of-function experiments: Abrogation of lateral inhibition leads to precocious commitment to the endocrine lineage; Lateral stabilization causes biphasic expression of pro-exocrine factors; And cell type ratios can be modulated by up/downregulation of the strength of lateral stabilization. Moreover, the coupling of the two feedback mechanisms causes a multi-stability of spatial patterning solutions that, in the presence of noise, generates a scattered distribution of endocrine cells, as observed in the central part of the pancreatic epithelium. Our analysis shows that the scattering pattern arises as a side-effect of noise on the relative timing of the two feedback mechanisms. This enables two qualitatively different patterns to appear concomitantly and persist side-by-side. The scattering pattern predicted by our model is characterized by the presence of small clusters of endocrine cells in local checkerboard-like patches, rather than a uniform random distribution. Although it remains unclear whether pro-endocrine  $Ng3^+$  cells actually appear as clusters within the primitive pancreatic epithelium, it is known they rapidly aggregate after their delamination from the epithelium. It may thus be speculated that the specification of endocrine cells in a clustered fashion serves as a pre-pattern to expedite the formation of aggregates that eventually form the islets of Langerhans.

Unlike for lateral inhibition, the molecular pathways underlying lateral stabilization remain unknown. Yet, based on the mathematical conditions under which the multistability occurs, receptor-ligand interactions can be excluded. Rather, our model predicts that lateral stabilization depends on homotypic binding of membrane-bound proteins. Recent studies offer several possible candidates. For instance, Minami et al.<sup>162</sup> demonstrated that disruption of cadherin-mediated cell-cell adhesion is required for the induction of dedifferentiation of adult acinar cells *in vitro*. This suggests that cadherins may be involved in stabilizing exocrine factors and maintaining acinar cell fate. Another, potentially complementary, possibility is that gap junctional communication mediates acinar cell fate stability. This is supported by experimental evidence that inhibition of *Mist1* activity, a key regulator of gap junctional communication in exocrine cells and downstream target of *Ptf1a*, also causes the dedifferentiation of acinar cells<sup>161,188</sup>.

Independent of the precise molecular realization, the fact that the both inhibitory and stabilizing mechanisms of contact-mediated induction have been reported for a wide range of tissues suggest that the model and results presented here for the pancreas may also be applicable to other developing tissues.

---

## TRANSDIFFERENTIATION OF PANCREATIC CELLS BY LOSS OF CONTACT-MEDIATED SIGNALING\*

---

### 6.1 INTRODUCTION

In the course of embryonic development, cells become progressively more specialized. Yet, it is becoming increasingly clear that adult differentiated cells retain the ability to change cell fate under certain conditions<sup>219,220</sup>. Novel approaches in regenerative medicine aim at harnessing this cell type plasticity in order to replace diseased or damaged tissue by targeted conversion of cells from other tissues<sup>221</sup>. Transdifferentiation, also known as lineage conversion, from one cell type to another often involves a dedifferentiation step to reinstate multipotency, but it is also possible to force cells to switch lineages directly<sup>222</sup>. Cells can be reprogrammed by ectopic expression of specific transcription factors using viral transduction<sup>223,224</sup>. However, some cell types can also be converted without genetic manipulation, by merely changing the cellular microenvironment. For many purposes, microenvironment-induced conversion may be preferable since it avoids the risks of random viral integration<sup>225</sup>. Contact-mediated signals from neighboring cells constitute a major part of the cellular microenvironment and recent studies have highlighted the importance of cell-cell contacts and surface-bound signals for pluripotency and cell type stability<sup>131,162,188,226–229</sup>. Yet, little is known about the regulatory effects of contact-mediated signals on cell fate stability and cell type conversion. In this paper, we investigate the role of contact-mediated signaling mechanisms in transdifferentiation by a theoretical study of cell fate control in the pancreas.

The pancreas is an organ with dual exocrine/endocrine functions. Acinar cells produce digestive enzymes that enter into the gut, whereas  $\alpha$  and  $\beta$ -cells, organized in the islets of Langerhans, release hormones into the blood stream for glucose homeosta-

---

\* This chapter is based on the publication: Walter de Back, Roland Zimm and Lutz Brusch, *BMC Systems Biology*, 7:77, 2013. Author contribution: Walter de Back conceived the study and constructed the model. Simulation, data analysis and bifurcation analysis were performed together with Roland Zimm. Walter de Back wrote the paper.

sis. Disruption of this homeostasis in diabetic patients is caused by a loss of functional  $\beta$ -cells. Conversion of cells from other pancreatic tissues into new  $\beta$ -cells has been proposed as a replacement therapy<sup>122</sup>. Acinar cells are interesting candidates as a source for transdifferentiation because of the common developmental origin of exocrine and endocrine cells as well as the abundance of acinar cells in the pancreas<sup>187</sup>. In fact, reprogramming of acinar cells into new  $\beta$ -cells has already been demonstrated *in vivo* in mice using ectopic expression of key transcription factors using viral transduction<sup>125</sup>. Intriguingly, such transdifferentiation has also been demonstrated in *in vitro* cultures without genetic manipulation, using only microenvironmental changes<sup>152,230–232</sup>. These studies show that adult acinar cells spontaneously dedifferentiate upon loss of cell-cell contacts by enzymatic tissue dissociation. Transcription factors and signaling pathways such as Notch signaling are reactivated which normally are only expressed during development. These progenitor-like cells can be converted into  $\beta$ -cells, although the yield is typically very low<sup>152,231–233</sup>. Interestingly, it has been found that the efficiency of lineage conversion can be improved dramatically by inactivation of Notch signaling<sup>154</sup>.

These findings suggest that at least two contact-mediated or lateral signaling pathways are involved in acinar-to- $\beta$ -cell conversion. First, dedifferentiation seems to be controlled by the loss of a stabilizing signal that is mediated by contact with adjacent acinar cells and is required for the maintenance of the acinar identity<sup>187</sup>. Second, redifferentiation into the endocrine lineage of islet cells seems to be hampered by contact-mediated Notch signaling<sup>154</sup> in a mechanism known as lateral inhibition, as previously described for pancreas development<sup>147</sup>. Understanding how these lateral signaling pathways act together in regulation of cell type stability and conversion dynamics can be an important step towards the development of non-genetic methods of  $\beta$ -cell neogenesis.

In this study, we construct and analyze a mathematical model that combines gene regulation with two contact-mediated signaling mechanisms: lateral inhibition and lateral stabilization. Using a combination of bifurcation analysis and numerical simulation, we find that multistability of gene expression states underlies the potential of acinar-to-islet cell conversion. Whereas loss of lateral stabilization causes a step-wise conversion through a multipotent progenitor-like state, additional loss of lateral inhibition induces the direct transdifferentiation from acinar to islet cells. In addition, cell density as well as the size and structure of cellular aggregates are found to affect the efficiency of conversion. Our results demonstrate that the combination of two lateral signaling mechanisms suffices to reproduce observations of acinar-to-islet cell conversion. By clarifying the role of lateral signals in lineage conversion, this new theoretical framework may contribute to improving strategies of microenvironment-induced transdifferentiation in general and to  $\beta$ -cell neogenesis in particular.

## 6.2 METHODS

Gene regulatory networks can be mathematically modeled and analyzed in terms of differential equations<sup>189</sup>. This can help to understand the complex feedback mechanisms underlying cell fate control<sup>137</sup>. By means of model analysis, one can reveal the existence of attractors that represent cellular phenotypes and understand the dynamics between states<sup>136,234,235</sup>. Using such a systems biological approach, we have previously shown that the results of genetic reprogramming experiments in the pancreas can be predicted from the hierarchical topology of the underlying gene regulatory network<sup>146</sup>. In the present study, we construct a minimal model of the gene regulatory network and contact-mediated signaling pathways underlying endocrine/exocrine cell fate decisions and maintenance in the pancreas and analyse this model using a combination of bifurcation analysis and tissue-scale lattice simulation.

The state of each cell is specified by four variables, A, X, Y, Z representing the expression levels of key transcription factors. Whereas X and Y correspond to core fate-determining genes and are involved in contact-mediated signaling, the factors A and Z represent up- and downstream factors (see figure 6.1). More specifically, the factor X represents the pro-endocrine transcription factor *Ngn3* that is transiently expressed during early pancreas development and participates in Notch-mediated lateral inhibition<sup>147,173</sup>. *Ngn3* activates the expression of the membrane-bound Notch ligand *Delta-like1* (*Dll1*)<sup>147</sup>. Reversely, activated Notch signaling causes inhibition of *Ngn3* by the transcriptional repressor *Hes1*<sup>148</sup>. As a result, neighboring cells compete for endocrine commitment by mutual inhibition of *Ngn3* expression, in a mechanism called lateral inhibition<sup>149,151</sup>. The factor Z represents a terminal endocrine fate marker downstream of *Ngn3* such as *Isl1*<sup>173,236</sup> that, once induced, retains its expression by positive auto-activation. As an islet cell maturation factor, it acts to repress the expression of upstream factor A.

The factor Y is interpreted as *Ptf1a*, which is the only transcription factor known to be necessary and sufficient to induce the exocrine cell fate<sup>174,176</sup>, but is expressed in all pancreatic progenitor cells<sup>175</sup>. Based on experimental evidence that adult acinar cells lose *Ptf1a* expression upon loss of physical cell-cell contact<sup>152,187,230–232</sup>, we assume that factor Y is involved in lateral stabilization. Lateral stabilization provides a positive feedback loop between Y-expressing neighboring cells (see chapter 5). The rate of Y production is up-regulated by its simultaneous expression in neighboring cells. Mathematically, this is represented by a multiplication, such that non-Y-expressing cells do not participate in lateral stabilization. Although the molecular details of a lateral stabilization pathway are unclear, such conditional activation is, in principle, consistent with both cadherin/ $\beta$ -catenin signaling<sup>162</sup> as well as with *Mist1*-mediated gap junctional communication<sup>188</sup>. In both cases, cells need to express monomeric proteins that form homotypic transmembrane complexes in order to signal to adjacent cells.



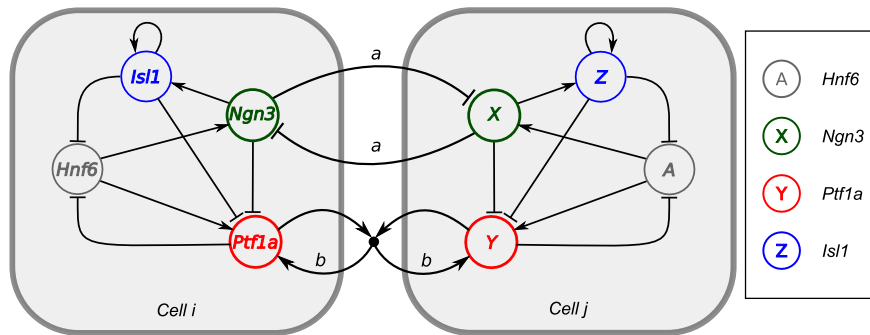


Figure 6.1.: **Gene regulation and lateral signaling network.** In cell i, the common names of the transcription factors are used. In cell j, these are replaced by the respective model variables. Cells i and j are coupled by lateral inhibition of factors X, and by lateral stabilization between factors Y. For each cell, the upstream factor A induces expression of X and Y, while X also induces Z expression, which activates itself. Both endocrine factors X and Z antagonize exocrine factor Y. Once differentiated, the markers Y and Z down-regulate A. Parameters in small lower case represent strengths of the interactions.

Although the endocrine and exocrine markers are mutually exclusive<sup>194</sup>, the underlying regulatory mechanisms remain unresolved. One model holds that *Nkx6.1*, a pro-endocrine factor downstream of *Ngn3*<sup>195,196</sup>, antagonizes the expression of *Ptf1a*<sup>197</sup>. Independent of the precise molecular pathway, we assume that (pro-)endocrine factors X and Z independently suppress the expression of Y leading to the restriction of the latter factor to the exocrine compartment.

Both *Ngn3* and *Ptf1a* are known to be induced by the upstream factor *Hnf6*, either directly<sup>190</sup> or indirectly<sup>191,192</sup>. To reflect this fact in the model, factor A induces the expression of X and Y. Both *Hnf6* and *Ngn3* are down-regulated during late developmental stages and are not expressed in the adult pancreas under normal circumstances<sup>173</sup>. In the model, this is captured by negative feedback of the terminal islet and acinar markers, Z and Y, on the inducing factor A. Indirectly, this also causes the down-regulation of X.

These gene-gene and cell-cell interactions can be formulated in terms of the following system of stochastic differential equations using Hill kinetics (parameters as in table 6.1):

	Symbol	Description	Value	
			<i>embryo</i>	<i>adult</i>
<b>Variables</b>	A	Expression of transcription factor <i>Hnf6</i>	1	0
	X	Expression of transcription factor <i>Ngn3</i>	0	0
	Y	Expression of transcription factor <i>Ptf1a</i>	0	1
	Z	Expression of transcription factor <i>Isl1</i>	0	0
	$\bar{X}$	Average <i>Ngn3</i> expression in neighboring cells	0	0
	$\bar{Y}$	Average <i>Ptf1a</i> expression in neighboring cells	0	1
<b>Parameters</b>	a	Strength of lateral inhibition $X \vdash \neg X$	1000	
	b	Strength of lateral stabilization $Y \leftrightarrow Y$	2000	
	c	Strength of inhibition $X \vdash Y$ and $Z \vdash Y$	500	
	q	Strength of induction $A \rightarrow X$ and $A \rightarrow Y$	$10^{-4}$	
	r	Strength of inhibition $Y \vdash A$ and $Z \vdash A$	100	
	s	Strength of autoactivation $Z \rightarrow Z$	50	
	n	Hill coefficient, nonlinearity of reactions	3	
	$\eta_x$	Noise amplitude on X	$10^{-3}$	
	$\eta_y$	Noise amplitude on Y	$10^{-3}$	

Table 6.1.: **Variable and parameter values of the mathematical model** (equation 6.1-6.4). The values of the variables refer to the embryonic and adult acinar initial condition.

$$\frac{dA}{dt} = \frac{1}{1 + rY^n + rZ^n} - A \quad (6.1)$$

$$\frac{dX}{dt} = \frac{qA^n}{q + a\bar{X}^n} - X + \xi_x(t) \quad (6.2)$$

$$\frac{dY}{dt} = \frac{qA^n + b(Y\bar{Y})^n}{q + b(Y\bar{Y})^n + cX^n + cZ^n} - Y + \xi_y(t) \quad (6.3)$$

$$\frac{dZ}{dt} = \frac{X^n + sZ^n}{1 + sZ^n} - Z \quad (6.4)$$

The terms  $\bar{X}$  and  $\bar{Y}$  denote the average expression of X and Y in the directly adjacent neighboring cells. To implement lateral inhibition, production of X is inhibited by the expression of X in neighboring cells,  $a\bar{X}^n$ , independent of its own activation. In contrast, the multiplicative term representing lateral stabilization,  $b(Y\bar{Y})^n$ , acts to stabilize a pre-existing expression. This requires the cell-autonomous activity of Y in both cells.

The additive stochastic terms  $\xi(t)$ , accounting for variability in gene expression or signaling noise, are random variables with a Gaussian white noise distribution  $N(0, \eta)$

with mean 0 and amplitude  $\eta$ . The Hill coefficient  $n$  is chosen such that the system exhibits non-linear step-like behavior ( $n = 3$ ). The model variables are scaled in such a way that the steady state expression of all factors is between 0 and 1. Parameter values are chosen such that the acinar cell fate (cells with  $Y \approx 1$ ) and islet cell fate ( $Z \approx 1$ ) are mutually exclusive. For brevity, in the presentation of the results below,  $Y^+$  cells are acinar,  $X^+$  cells are islet progenitors, and  $Z^+$  cells are islet cells.

The states and (in)stabilities of the above model were studied using bifurcation analysis. Numerical simulation of a hexagonal lattice of cells was performed to study the spatiotemporal dynamics at the tissue scale. Analysis and numerical simulation were performed using GRIND (phase plane analysis)<sup>200</sup>, XPPAUT (bifurcation analysis)<sup>201</sup> and our modeling environment Morpheus (lattice simulations)<sup>55</sup> (see chapter 3). The stochastic differential equations were solved using the 2<sup>nd</sup> order Heun-Maruyama method with time step size  $dt = 0.02$ . The simulation model is available in the Supplementary Online Material under <http://walter.deback.net/thesis>.

## 6.3 RESULTS

### 6.3.1 Multistability of acinar and islet cell fates

Cell fates are characterized by stable patterns of gene expression. Whether a set of interacting genes is able to reach one or more stable states depends on their interaction topology as well as on the strengths of interaction. To investigate the cell fates that can appear in our model, we studied the existence of stable states and their dependence on parameter values for lateral signaling by performing a bifurcation analysis.

Due to lateral signaling, the fates of individual cells depend on the states of neighboring cells. Therefore, we analyzed a system of three cells representing a minimal tissue that is able to show all possible configurations present in larger systems (the mixed state does not occur for less than 3 cells). To study how the stability of cell fates changes while varying the strength of the lateral stabilization mechanism  $b$ , we recorded the summed expression level of exocrine factor  $Y$ . This reduces the high-dimensional state space to a single dimension and provides information on cell fates as well as their spatial pattern. The solid lines in the bifurcation diagram in figure 6.2A show that  $Y$  expression has three stable states over a wide range of parameter values. For these values of  $b$ , the three cells can have either acinar fates ( $Y = 3$ ), islet cell fates ( $Y = 0$ ) or have mixed fates ( $Y = 2$ ), depending on initial conditions or history of gene expression.

This multistability of acinar and islet cell fates has several interesting consequences. The key observation is that a critical value  $b_c$  exists, below which the stable steady state for the acinar fate disappears, while the islet cell fate remains stable. Thus, loss of the stabilizing effect of lateral signaling effectively moves the system towards a region in parameter space where the acinar cell fate no longer exists. Therefore, upon

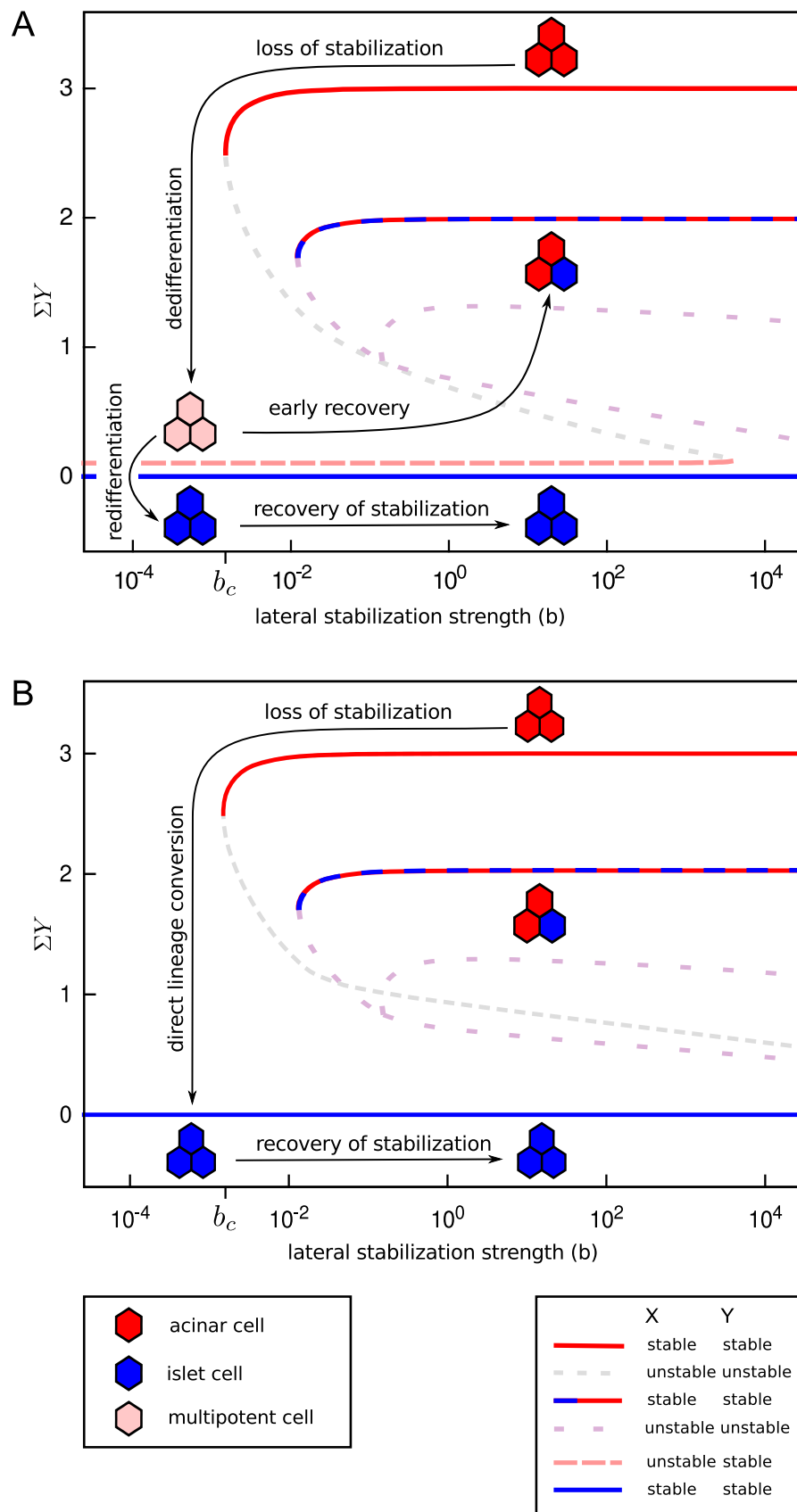


Figure 6.2.: **Bifurcation analysis of three-cell system.** (Caption on next page)

Figure 6.2.: Stability of cell fates change as a function of strength of lateral stabilization. Bifurcation diagram showing stable attractors (solid) and unstable states (dashed) for a minimal tissue consisting of three-cells (hexagons). Arrows indicate trajectories after loss of stabilization. (A) In presence of lateral inhibition,  $a > a_c$ , loss of stabilization,  $b < b_c$ , causes dedifferentiation towards a progenitor-like multipotent state. If lateral stabilization is recovered at this early stage, the developmental process is recapitulated and a mixed pattern of both cell fates arises. In contrast, if stabilization remains inhibited, cells redifferentiate into islet cells. (B) In absence of lateral inhibition,  $a < a_c$ , loss of stabilization results in direct lineage conversion, due to the absence of a progenitor-like multipotent state. This unstable multipotent state vanishes at  $a_c$  in a saddle-node bifurcation with another solution branch of similar  $Y$ -values but higher  $Z$  activity which is additionally unstable against perturbations in  $Z$  and therefore omitted in (A). Note that  $\Sigma Y$  is a projection of a high (12)-dimensional space, such that intersections do not imply bifurcations or changes in stability as these need not intersect in the actual state space. In the legend, the stability of  $X$  or  $Y$  means (un)stable with respect to perturbations in variable  $X$  and  $Y$ . With parameters as in table 1,  $a_c \approx 0.0017$  and  $b_c \approx 0.012$ .

such a change in parameter values, acinar cells lose their exocrine markers and dedifferentiate spontaneously. In the presence of lateral inhibition (fig. 6.2A) cells adopt a multipotent progenitor-like fate. This state is stable against perturbations in  $Y$ , but unstable against perturbations in  $X$ , which implies that noise on  $X$  can change this state. If lateral stabilization is recovered at this multipotent stage, the system moves towards a steady state with mixed acinar and islet cell fates, recapitulating the cell fate decision and spatial pattern observed during pancreas development<sup>101</sup> (see chapter 5). If, however, disruption of lateral stabilization continues, cells differentiate into the islet cell lineage. After completing the lineage conversion, the islet fate is stable in the sense that recovery of lateral stabilization does not reverse conversion.

Interestingly, the bifurcation analysis shows a different behavior in the absence of lateral inhibition (fig. 6.2B). In this case, multipotent progenitor-like steady state does not exist. This implies that acinar cells cannot dedifferentiate towards a progenitor-like state upon loss of lateral stabilization. Instead, cells undergo direct lineage conversion from the acinar to the islet lineage, rather than passing through a state of multipotency.

In conclusion, bifurcation analysis reveals (1) that lateral stabilization accommodates multistability of the acinar and islet cell states, (2) that transient loss of lateral stabilization can cause the conversion of acinar to islet cells and (3) that concomitant suppression of lateral inhibition leads to direct conversion, bypassing the multipotent progenitor-like state. Yet, bifurcation analysis does not provide insight into the spatiotemporal dynamics for which we next turn to numerical simulations.

### 6.3.2 Cell fate decision and patterning during pancreas development

Cells undergoing acinar-to-islet cell conversion transiently express various factors and activate signaling pathways normally only observed during development (*Pdx1*, *Hnf6*, *Ngng3*, *Notch*, *Dll1*)<sup>154,232</sup>. This suggests that at least a part of the developmental regulatory network is reactivated<sup>187</sup> and that cell fate decisions during organogenesis and cell type conversion are governed by the same regulatory mechanisms. Under this assumption, the model proposed here for acinar-to-islet cell conversion is expected to reproduce the cell fate decision between the exocrine or endocrine lineage during embryonic development.

To test whether the proposed model holds for the conditions during embryonic development, we simulated the model using initial conditions that represent the gene expression in early pancreatic progenitor cells. In the mouse, the inductive factors *Hnf6* and *Hnf1β*, that act upstream of lineage-associated factors *Ngng3* and *Ptf1a*, are first detected around E9<sup>175</sup>. At this stage, *Ngng3* and *Ptf1a* themselves are not yet expressed. Accordingly, the early embryonic state is accounted for in our model by the homogeneous expression of A (table 1).

Figure 6.3B shows that during simulation, A activates the expression of both X and Y. For a transient period, these factors are co-expressed in all cells at low or intermediate levels of expression. The “promiscuous” co-expression is typical of multipotent progenitor cells and is also observed in pancreatic progenitors<sup>194</sup>. During this phase, mutual inhibition between cells maintains a low-level expression and thereby suppresses differentiation into either lineage, similar to the role of Notch signaling in pancreatic development known as “suppressive maintenance”<sup>210</sup>. After noise introduces variation in X expression between cells, these differences become amplified by lateral inhibition and result in a divergence of X expression. Factor X activates islet cell differentiation by activating Z and is only transiently expressed itself, as is known for *Ngng3*. Reversely, in the  $X^-$  surrounding cells, Y is no longer inhibited and is upregulated. Through lateral stabilization,  $Y^+$  cells induce the expression of Y in neighboring progenitor cells (with low Y expression) which results in wave propagation, in a process traditionally known as homeogenetic induction<sup>215</sup>. Maturation into either lineage results in suppression of upstream factor A which leads to the downregulation of the pro-endocrine factor X, while Y is maintained by lateral stabilization. In line with experimental observations, both factors (*Hnf6* and *Ngng3*) are not expressed after the cell fate decision and in the adult pancreas.

Interestingly, the spatial patterns generated by the model are also in line with reports of the scattered distribution of nascent islet cells in the early pancreatic epithelium<sup>177</sup>. The combination of lateral inhibition (creating an alternating pattern of acinar and islet cells) with lateral stabilization (creating homogeneous fields of acinar cells) results in the establishment of a scattered spatial distribution of endocrine cells in a mainly exocrine tissue (see figure 6.3C)<sup>101</sup>.

In short, under initial conditions representing early pancreas development, the key features of gene expression and patterning in the developing pancreas are reproduced by the model: (1) promiscuous expression of the lineage-associated factors *Ngng3* and *Ptf1a*, (2) the transient expression of the pro-endocrine factor *Ngng3* and (3) the scattered spatial patterning of committed islet cells.

### 6.3.3 Loss of lateral stabilization causes sequential conversion

To understand the dynamics of acinar-to-islet cell conversion upon loss of lateral stabilization, simulated cells were initialized with an acinar-like gene expression profile in which only the exocrine factor *Y* is expressed (see table 1). The system was initialized with lateral stabilization strength  $b > b_c$  to ensure the stability of the acinar-like state under these conditions. After a given period, lateral stabilization was lost,  $b = 0$ , marking  $t = 0$ .

As shown in figure 6.3B', the acinar state is stable as long as lateral stabilization strength  $b > b_c$ , representing intact acinar tissue. However, immediately following the loss of lateral stabilization, cells lose the expression of exocrine marker *Y*. The lack of the maturation factor *Y* leads to the re-activation of the upstream factor *A*. Since *A* induces low levels of both *X* and *Y*, at this stage, the expression pattern is identical to the early embryonic situation. Thus, loss of lateral stabilization causes cells to return towards the multipotent progenitor-like cell state. If the absence of lateral stabilization continues, the subsequent dynamics differ from the embryonic cell fate decision discussed above. Specifically, nascent islet cells arise in an alternating spatial pattern as a result of lateral inhibition between  $X^+$  cells (fig. 6.3C). Yet, this pattern is not stable. After a cell has committed to the islet lineage by transactivating the endocrine marker *Z*, it loses expression of *X*. Therefore, cells adjacent to endocrine  $Z^+$  cells are no longer inhibited and will start to express *X* themselves. As a result, some of the neighboring cells also commit to the endocrine lineage, after which the process is repeated. This step-wise conversion of cells within the tissue results in a complex spatiotemporal patterning process (fig. 6.3C'). Under these idealized conditions, eventually all cells commit to the islet cell lineage. If, however, lateral stabilization is recovered before cells have redifferentiated, the cell type conversion is arrested which significantly decreases the efficiency of conversion (data not shown). Recovery does not revert newly committed islet cells back to acinar fate, since the islet cell state is stable, independent of lateral stabilization.

These results are in line with *in vitro* experiments showing spontaneous dedifferentiation upon enzymatic disassociation and disruption of cadherin-mediated cell-cell adhesion<sup>152,162,232,233</sup>. Furthermore, these results suggest that acinar-to-islet cell conversion ensuing loss of lateral stabilization is a relatively slow process due to the fact that lateral inhibition prevents neighboring cells from completing transdifferentiation simultaneously.

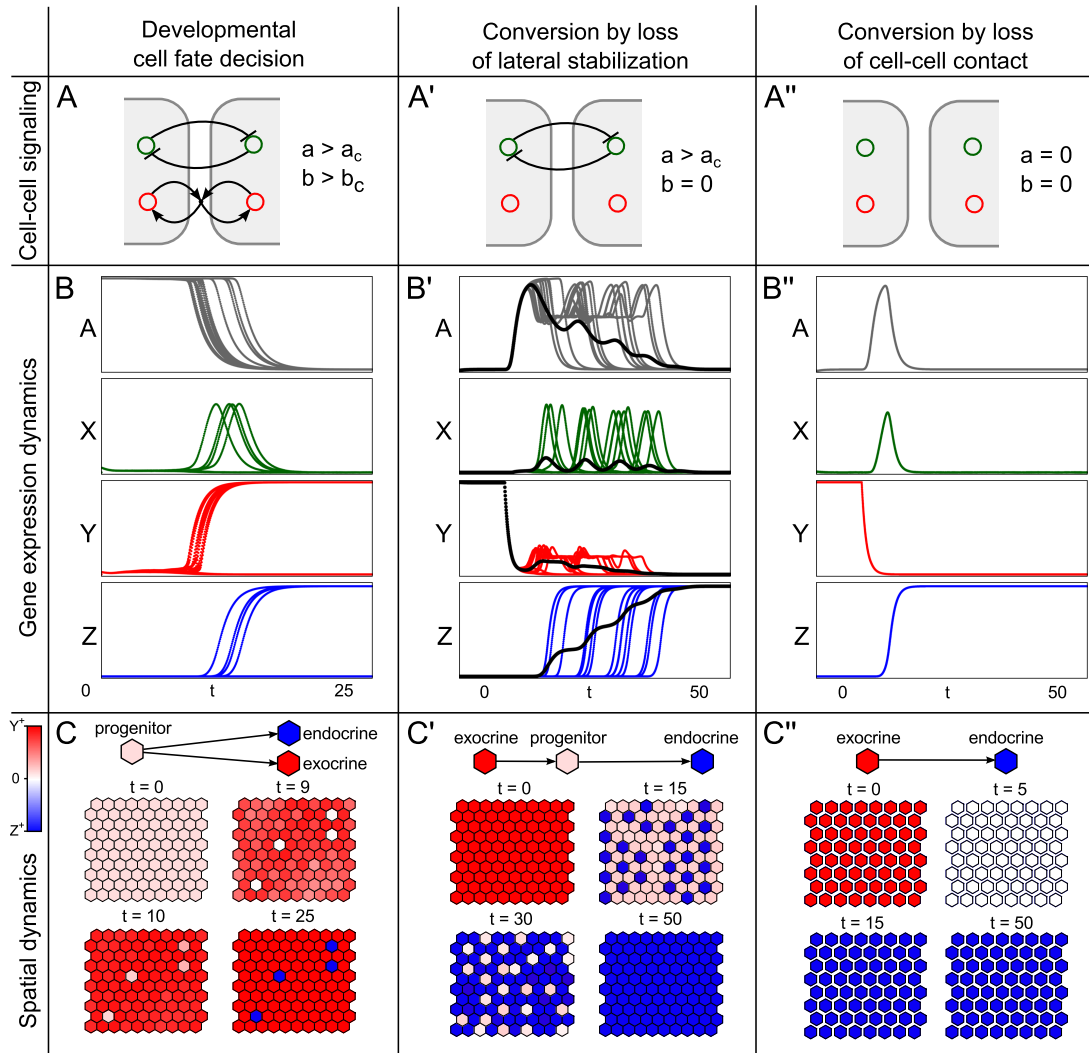


Figure 6.3.: **Dynamics of cell fate control.** Dynamics of cell fate decisions during development (left column) and lineage conversion (middle and right columns). (A) Sketch of cell-cell signaling condition. (B) Expression of transcription factors over time. A: *Hnf6*, X: *Ngtn3*, Y: *Ptf1a*, Z: *Isl1*. Black lines in B' depict population averages. (C) Emergent spatial patterns, representing cell fates by colors. Color coding:  $Y^+$  acinar cells are red,  $Z^+$  islet cells are blue, and  $Y^-Z^-$  cells are white. Initial condition for development is  $A = 1, X = Y = Z = 0$  and for conversion is  $Y = 1, A = X = Z = 0$ . Parameters as in table 6.1. Videos and simulation models are all three conditions are available in the Supplementary Online Material under <http://walter.deback.net/thesis>.



#### 6.3.4 Loss of lateral inhibition accelerates conversion

In the embryo, disruption of the Notch signaling pathway is known to cause precocious endocrine commitment<sup>147</sup>. Moreover, its inhibition in adult acinar cells can dramatically increase the efficiency of acinar-to-islet cell type conversion<sup>154</sup>. Since one of the roles of Notch signaling in the developing pancreas is lateral inhibition, we examined the dynamics of the model after a sudden loss of lateral inhibition. As before, we used the acinar-like initial conditions (table 1), but now both lateral stabilization and lateral inhibition were lost,  $a = b = 0$ , after a given period.

Immediately ensuing this manipulation, Y expression rapidly decreases, causing the reactivation of A expression, as described before. However, in this case, the dedifferentiated cells do not return to a multipotent state with “promiscuous” co-expression. Instead, all cells simultaneously upregulate the pro-endocrine factor X since they are not inhibited by their neighbors (see figure 6.3B’). Finally, after the transactivation of Z by X, the factors A and X are suppressed again, leading to an adult islet fate in all cells. Compared to the loss of stabilization, the additional loss of lateral inhibition results in a much faster dynamical process of lineage conversion. In line with results obtained *in vitro*<sup>154</sup>, our model shows that concomitant inhibition of lateral inhibition accelerates acinar-to-islet conversion. Here, this observation is explained by the fact that, under disruption of lateral inhibition, the unstable steady state representing the multipotent progenitor state does not exist, as predicted by bifurcation analysis (figure 6.2B).

Note that disruption of lateral inhibition alone ( $a = 0$ ,  $b > b_c$ ) does not affect acinar cell stability, since the pro-endocrine factor X, which is involved in this feedback between cells, is not expressed in adult acinar cells. Therefore, without loss of lateral stabilization, cells maintain their acinar identity.

#### 6.3.5 Cell density affects conversion efficiency

If the disruption of contact-mediated signaling influences the efficiency of acinar-to-islet cell conversion, loss of physical contacts between cells could replace molecular manipulation. To study the effect of cell-cell contacts, we performed simulations with varying densities of acinar cells. As expected, it was found that conversion efficiency increases with decreasing cell density (see figure 6.4). For extreme cases, the reason behind this is evident. At high densities, most cells have many contacts with neighboring acinar cells and the stabilizing positive feedback prevents their dedifferentiation. Conversely, at low density, most cells are isolated and do not receive stabilizing (or inhibiting) cell-cell signals. Consequently, these cells can complete transdifferentiation. However, for more realistic intermediate cases in which cells are part of small aggregates, the situation becomes nontrivial. Here, the probability of cell conversion depends on both size and shape of the cellular aggregate. Although cells in larger clusters are generally more stable, this stability also depends on the spatial arrangement

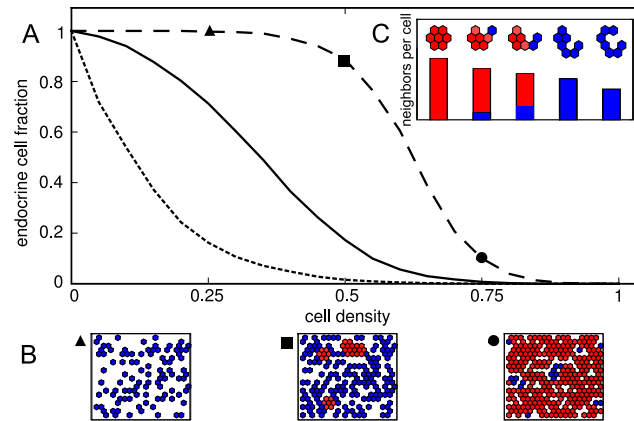


Figure 6.4.: **Cell density affects conversion efficiency** (A) The fraction of acinar cells that convert to islet cells increases with decreasing cell density, as shown for three values of lateral stabilization strength,  $b = 1$  (dotted),  $b = 0.1$  (solid),  $b = 0.01$  (dashed). (B) Examples of the steady-state situation (acinar cells in red and islet cells in blue) for three different cell densities as indicated on the dashed curve ( $b = 0.01$ , densities 0.25, 0.50 and 0.75). Note the presence of compact clusters of stable acinar cells in the middle panel. (C) Shape of cellular aggregates determines the efficiency of conversion. A decrease in compactness, measured as average neighbors per cell, increases the islet cell yield. Parameters as in table 6.1,  $b$  as indicated.

of cells in the aggregate (fig. 6.4C). Because the dedifferentiation of one cell weakens the stability of its neighboring cells, waves of dedifferentiation and conversion can propagate through the aggregate, depending on the average number of neighboring cells that reflects both density and configuration of a cell cluster.

These results show that, in the context of our model, the size and the structure of cellular aggregates affects the efficiency of lineage conversion. This implies that the degree of dissociation of acini by enzymatic digestion is predicted to have large impact on islet cell yield. More generally, the use of low cell densities or, alternatively, inhibition of reaggregation of cells, is predicted to increase the efficiency of acinar-to-islet cell conversion *in vitro*.

#### 6.4 DISCUSSION AND CONCLUSION

Forcing adult cells to change lineage by altering the microenvironment offers an alternative to the more risky method of virus-mediated nuclear reprogramming<sup>224,225</sup>. Apart from identifying of specific growth factors and small molecules that induce a particular lineage conversion *in vitro*, recent work in this direction also demonstrates that contact-mediated lateral signals are key regulators of cell fate maintenance and multipotency<sup>131,162,188,226–229</sup>. For instance, it was found that loss of cell-cell adhesion between adult acinar cells of the pancreas causes dedifferentiation and enables their

conversion into islet cells<sup>152,162</sup>. Together with more recent data showing that inhibition of contact-mediated Notch signaling between these cells significantly improves conversion efficiency<sup>154</sup>, this demonstrates that lateral signals are important regulators of cell fate control in the pancreas. However, despite the identification of a myriad of transcription factors and signaling molecules involved in lineage conversion, a coherent understanding of the roles of contact-mediated lateral signals in this process is lacking.

A systems biological approach can help to make sense of complex dynamic regulatory networks through the use of mathematical models and dynamical system theory<sup>137,189</sup>. In previous work, we have adopted this approach to construct a hierarchical multi-attractor model of the pancreatic transcriptional network to understand and propose nuclear reprogramming strategies<sup>146</sup>. In the present study, instead, we have focused on the role of contact-mediated signals on conversion dynamics to predict the outcomes of microenvironment-induced strategies for transdifferentiation.

We have presented a mathematical model that combines gene regulation and lateral signaling in pancreatic cells. We have demonstrated that the crosstalk of two contact-mediated signaling mechanisms (lateral inhibition and lateral stabilization) causes multistability in which both acinar and islet cell fates are stable. Our discovery of the multistable state explains why conversion of acinar to islet cells is possible, even without genetic manipulation. Inhibition of lateral stabilization destabilizes acinar cells and causes the dedifferentiation of acinar cells towards a progenitor-like multipotent state and invokes the subsequent step-wise conversion towards an islet cell fate. Moreover, we have shown that additional loss of lateral inhibition accelerates the conversion dynamics because, under these conditions, cells undergo a direct lineage switching, without passing through a multipotent state.

Altogether, our results provide a theoretical background to understand studies of acinar-to-islet cell conversion *in vitro*<sup>152,154,162,230,231</sup>. Moreover, this study offers several testable predictions, such as the impact of cell density, that may be used to improve the efficiency of micro-environment-induced conversion strategies. More generally, our results demonstrate that the crosstalk of multiple lateral signaling mechanisms can generate counterintuitive effects controlling cell fate stability as well as spatial patterning, which deserve further investigation. Furthermore, this study underscores that the identity of cells depends on the multicellular context of the tissue. Therefore, considering the feedback from the tissue level to the genetic level is important in order to understand how cell fate stability and plasticity are controlled.



Part III

MORPHOGENESIS

---

## INTRODUCTION

---

The processes that determine the shapes of the tissues of the organism are collectively called morphogenesis. In Part II, where we focused on the coupling of cell fate specification and spatial patterning, morphogenetic processes have been largely neglected. Here, as a second case study, the computational methods outlined in Part I are employed to investigate the establishment of tissue shape. Specifically, we study the formation of blood vessel networks by coupling cell-based models of motile chemotactic cells to reaction-diffusion systems representing signaling molecules and extracellular matrix components.

### 7.1 THE VASCULAR SYSTEM

Some small and simple multicellular organisms, such as hydra and planaria, depend on diffusion for the exchange of oxygen, nutrients and metabolites with the environment. However, in larger metazoan, the decreased surface to volume ratio causes diffusion to be an inefficient means of transport, since most cells do not have direct contact with the external environment. Circulatory systems, such as vascular networks, reduces the functional diffusion distance that nutrients, gases, and metabolic waste products must traverse through the transport of fluids<sup>237</sup>. The evolution of these systems have removed constraints on body size and geometry and is central to the emergence of larger multicellular organisms with complex body plans<sup>237</sup>.

The vascular system of vertebrates consists of three types of blood vessels: (1) arteries that carry blood away from the heart, (2) veins that transport blood back toward the heart and (3) capillaries. The capillaries are the smallest vessels of the body and are the ones that enable the actual exchange of material between the blood and the tissues. Capillary blood vessels are small tubes with a diameter of about 5-10 microns with walls composed of a single layer of endothelial cells. Diffusion and transcytosis through the endothelial cells that line the capillaries allows material exchange. To allow this exchange to cells and tissue throughout the body, capillaries form an extensive network. The total length of this network has been estimated to be around 100.000 km in a single adult human body<sup>238</sup>.

## 7.2 VASCULOGENESIS

Different cellular mechanisms are responsible for the formation and growth of blood vessels in the vertebrate embryo. The earliest vascular networks appear at an early stage during development after a morphogenetic process called vasculogenesis. Vasculogenesis takes place before the onset of blood flow and involves the differentiation and coalescence of isolated endothelial cell progenitors, angioblasts, from the mesoderm<sup>239</sup>. Vasculogenesis is the predominant mode of vascular development in the early embryo. Primitive embryonic vascular networks continue to grow by angiogenesis, an alternative mechanism of vessel growth based on invasive sprouting or division of pre-existing vessels. This is the main form of physiological and pathological mechanism of blood vessel growth and remodeling throughout adulthood, but naturally depends on the primitive vasculature assembled during embryonic vasculogenesis. Apart from the crucial role in the formation of blood vessels for the metabolic sustenance of developing organs, blood vessels are also a source for inductive signals for the development of organs such as the liver and the pancreas, even before the onset of blood flow<sup>163,240</sup>.

The first steps in the process of vasculogenesis involve the specification of endothelial cell progenitors from the mesenchymal cells of the mesoderm that, once formed, undergo a morphogenetic process in which they assemble into a network. During this process that takes place in a thin planar surface of the mesoderm, primary unicellular units associate into secondary multicellular vascular units that in turn fuse to give rise to a planar blood vessel network, called the primary capillary plexus. During the assembly, angioblasts differentiate into endothelial cells and form vascular lumen to allow blood flow.

Even before lumenization, the primary capillary plexus shows a striking characteristic pattern of a polygonal network. The signals and mechanisms that regulate the appearance of this network pattern remain unclear. Various possibilities have been proposed<sup>241,242</sup>. Possibly, the endodermal tissues provide a pre-pattern for the spatial organization of the vascular network, under strict genetic determination. This would require precise spatio-temporal control of the expression of provascular growth factors. Yet, this seems to be at odds with the substantial differences observed at the cellular scale between network in embryos of the same species. Transplantation experiments have also shown that early vascular network are adaptive and tissue grafts get rapidly integrated in their new host environment. Moreover, polygonal networks can also be formed by endothelial cells *in vitro*, where the presence of a pre-pattern can be ruled out. Rather than a genetic pre-pattern, this suggest that vascular pattern formation depends on a dynamic self-organized mechanism in which angioblasts coalesce into networks in response to cues from their local microenvironment.

### 7.3 MATHEMATICAL MODELS OF VASCULAR NETWORK FORMATION

Various mathematical models have been proposed that explore different mechanisms of self-organization that might underlie the formation of vascular network, in most cases supported by data on early stages of network formation in *in vitro* assays of Human Umbilical Vascular Endothelial Cells (HUVECs) in Matrigel. This work has resulted in a surprisingly large number of different mechanisms and mathematical models that allow for network formation (see<sup>241–244</sup> for reviews). Three types of models can be distinguished on the basis of their main biological assumptions: activator-inhibitor models<sup>72,245</sup>, mechanical models<sup>246–249</sup> and chemotactic models<sup>102,113,115,250–253</sup>.

#### 7.3.1 Activator-inhibitor models

Activator-inhibitor models are reaction-diffusion models in which the interaction between at least two substances, an autocatalytic activator species and an antagonistic inhibitor species, can result in spontaneous pattern formation provided that the diffusion of the inhibitor is fast compared to the diffusion of the activator<sup>44,70</sup>. In related models known as substrate-depletion models, the inhibitor species is replaced by the depletion of a substrate that is required for synthesis of the activator. Koch and Meinhardt showed that hierarchically coupling of a substrate-depletion model with an activator-inhibitor model is sufficient to obtain reticular network patterns<sup>245</sup>. While the primary substrate-depletion system generates a pattern of isolated peaks (lacunae), the secondary activator-inhibitor system produces stripes (vessels) around these peaks. This mechanism results in polygonal structures with a striking resemblance to primitive vascular networks (see figure 7.1A). Moreover, this mechanism shows size regulating properties keeping a constant characteristic polygons size, even in a growing domain such as the embryo<sup>245</sup>. Nevertheless, the molecular identities of the four interacting substances remain elusive.

#### 7.3.2 Mechanical models

An alternative hypothesis is that the formation of the vascular network is driven by the mechanical pulling activity of cells on the extracellular matrix (ECM)<sup>246–249</sup>. Mathematical models of the Murray-Oster-Harris type<sup>246</sup> can describe changes in cell density as a result of (1) passive motion due to the attachment of cells on a moving matrix and (2) a strain-biased random motion along regions of aligned matrix fibers. Analysis of the pattern-forming capabilities of such models show that network formation is possible by purely mechanical cell-matrix interactions (fig. 7.1B). Moreover, it demonstrates that network formation crucially depends on cell mechanical aspects such as the ability to create traction forces as well as the mechanical properties of the matrix such as its stiffness, in line with experimental observations<sup>246–249</sup>. Although in those studies, random



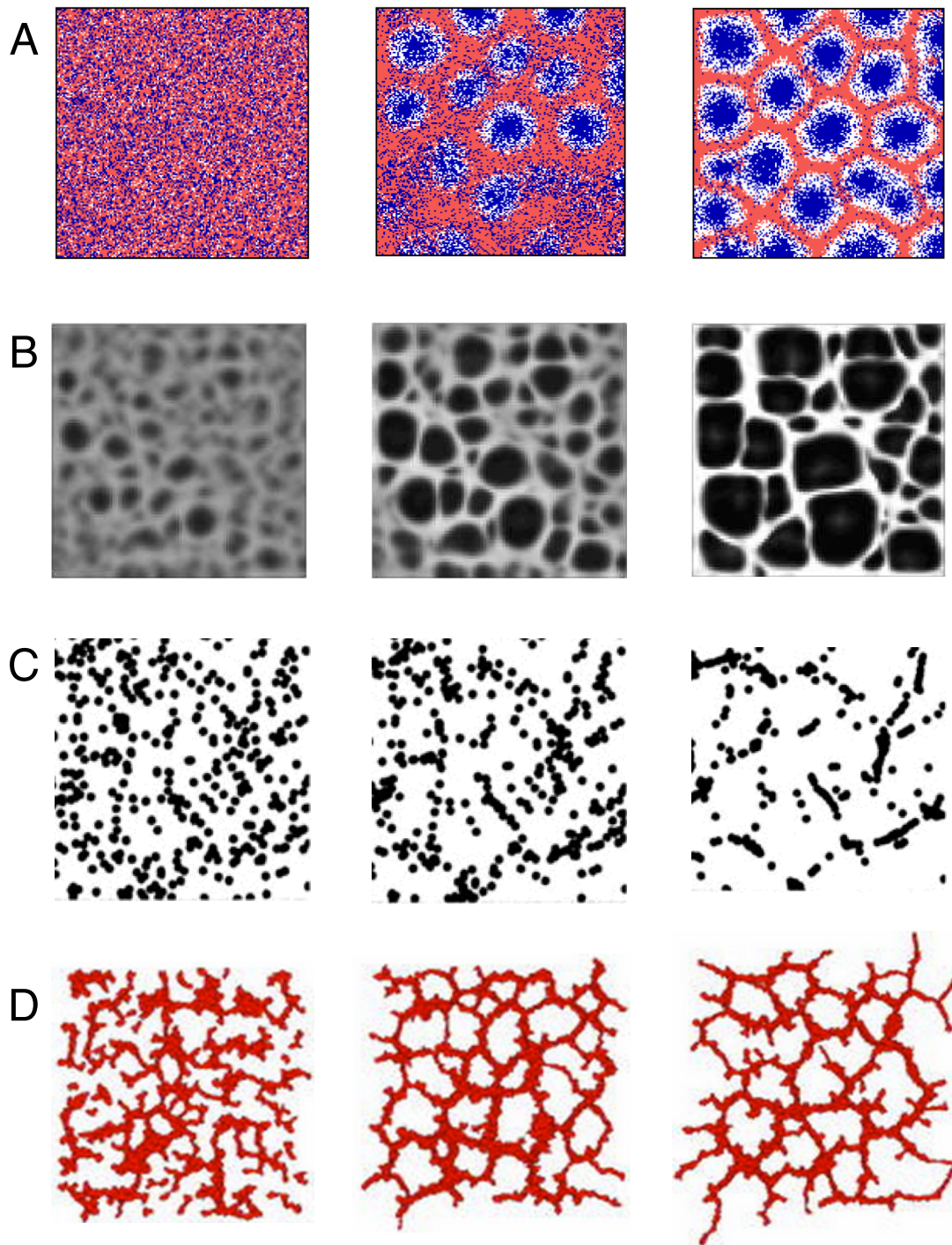


Figure 7.1.: Previous mathematical models of vascular network formation. (Caption on next page.)

Figure 7.1.: **Previous mathematical models of vascular network formation.** Three simulation time points of each model are shown. (A) Activator-inhibitor model. The primary substrate-depletion model creates spots (blue), while the secondary activator-inhibitor model creates polygonal networks (red)<sup>245</sup>. (B) Mechanical model. Traction forces exerted by cells on the extracellular matrix cause cellular networks of increased cell densities (greyscale intensity shows cell densities)<sup>248</sup>. (C) Continuum model of autocrine chemotaxis model. Networks arise through chemotaxis towards a chemoattractant produced by endothelial cells (black dots show cell centroids using a threshold on cell densities)<sup>250</sup>. (D) Hybrid continuum/discrete model of autocrine chemotaxis. Steep gradients of the chemoattractant, modeled in a continuous diffusion model causes cells, modeled by the cellular Potts model, to assemble into network patterns<sup>113,254</sup>.

cell motility was found to have little influence on network formation, a recent study found the opposite effect. Using a cellular Potts model coupled to a finite element model to model mechanical cell-matrix interactions, Merks and colleagues found that strain-biased cell motility was the driving force behind network formation<sup>249</sup>. These studies, and accumulating experimental data, show that mechanical tensions between cells and extracellular matrix are important in the vascular networks. However, mechanical interactions alone cannot explain the crucial role of chemical factors such as VEGF (vascular endothelial growth factors) known to be essential for vascular network formation in the embryo as well as *in vitro*.

### 7.3.3 Chemotaxis models

Chemotaxis models, the third type of models, attempt to explain network formation as a result of cell motility in the direction of a gradient of a diffusive substance, a chemoattractant. Most mathematical models of chemotaxis assume the release of a diffusive chemoattractant, identified as VEGF, by endothelial cells<sup>113,250–253</sup>. Using a continuum approach describing changes in cell density and spatial distribution, these models<sup>250,251</sup> demonstrate the possibility of network formation through the aggregation of cell densities along polygonal structures, given directional persistence of cells and sufficiently steep gradients of the chemoattractant (fig. 7.1C). These morphologies are, however, transient and eventually stabilize into disconnected structures and, moreover, fail to reproduce temporally correct network coarsening. In an approach to correct for these observations, the chemotaxis models was reformulated using a cellular Potts model model that allows explicitly accounting for cell shape, while maintaining the biological assumptions underlying the chemotaxis model<sup>113,252,253,255</sup>. With those models, it was shown that a number of additional mechanisms at the cellular level, including cell adhesion<sup>255</sup>, active elongation<sup>253</sup> and contact-inhibition<sup>113</sup> can, in con-

junction with autocrine secretion of a chemoattractant, suffice to reproduce vascular network and subsequent remodeling (fig. 7.1D).

#### 7.4 PARACRINE MODEL OF VASCULAR NETWORK FORMATION

In the following chapters, two studies are described in which an alternative mathematical model based on paracrine chemotaxis is proposed.

##### 7.4.1 Paracrine chemotaxis model with extracellular matrix retention

The chemotaxis models describe above show a robust ability to form network patterns under varying conditions. However, they are based on some hypothesis that are not fully substantiated by biological evidence. For instance, the diffusivity of the chemoattractant is required to be orders of magnitude lower than that of VEGF, in order to establish the steep gradients that are needed for network formation. More importantly, they rely on the assumption that endothelial cells release VEGF, acting as an autocrine chemoattractant (see figure 7.2A). Despite the fact that endothelial cells, including angioblasts and HUVECs, express VEGF receptors (VEGFR2, also known as Flk1), it is unclear whether endothelial cells produce and release their ligand. Instead, there are several indications that suggest an alternative source of VEGF. It is known that, in the embryo, the endodermal tissues near the nascent vascular plexus are a potent source of proangiogenic signals, including VEGF<sup>256,257</sup>. Moreover, in the commonly used HUVECs *in vitro* assays, the administration of external VEGF is required for network formation and usually provided as part of the cell growth medium (EGM2)<sup>258</sup>, indicating that the VEGF produced by cells, if any, is not sufficient. It is therefore likely that endothelial cells themselves are not the source of VEGF. By consequence, VEGF does not act as an autocrine, but rather as a paracrine chemoattractant.

This, however, raises an important new question: if VEGF arrives to the nascent vascular plexus by diffusion from an external source, how can it encode for fine-grained reticular networks? As mentioned, chemotaxis models depend on the existence of steep chemotactic gradients to avoid the formation of clusters rather than networks. How, if not by autocrine release, can such steep gradients be established?

The answer may come from the extracellular matrix. Apart from the mechanical role of the matrix discussed above, it can have a role in mediating and retaining signaling factors<sup>259–261</sup>. Interestingly, matrix molecules such as fibronectin and heparan sulphates are rich in VEGF binding domains. Moreover, it is plausible that these matrix molecules are either produced or modified by endothelial cells or are accumulated around endothelial cells through mechanical pulling. It could therefore be possible to generate the steep gradients required for network formation under these assumptions that paracrine VEGF binds to matrix molecules in pericellular regions (see figure 7.2B).

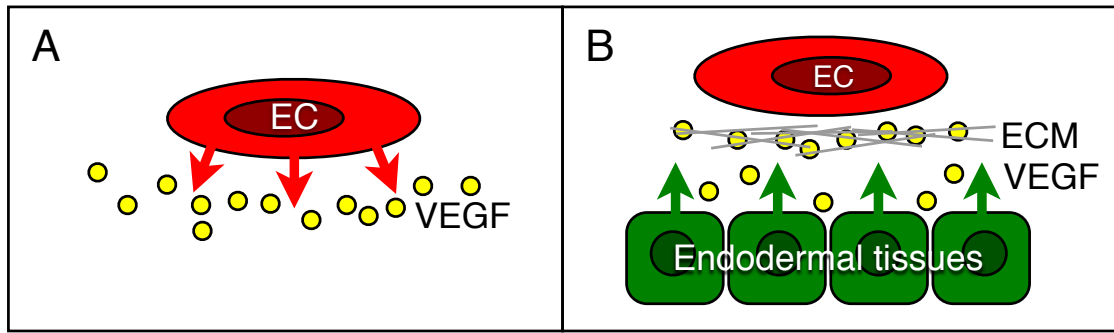


Figure 7.2.: **Autocrine and paracrine models of vascular network formations.** The key difference is in the assumed source of VEGF. (A) In the autocrine models, endothelial cells (EC) produce their own chemoattractant. (B) In the paracrine models, the chemoattractant is produced in nearby endodermal tissues. Spatially restricted guidance cues in pericellular regions can arise by binding of VEGF to extracellular matrix (ECM).

To investigate the possibility of pattern formation in this paracrine chemotaxis models, I collaborated with Alvaro Köhn-Luque to construct a computational model based on previous work by Merks and colleagues<sup>113,253,254</sup>, described in chapter 8. This model is an adaptation of the previous models in which we excluded the additional cellular mechanisms such as adhesion and contact-inhibition and included reaction-diffusion equations describing the key components of the paracrine chemotaxis model: (1) external homogeneous production of VEGF, (2) production or modification of ECM by endothelial cells, (3) binding of VEGF to matrix components, and (4) chemotaxis of cells towards bound rather than diffusive VEGF. As the results of this study show, this model is able to produce vascular network formation with similar morphogenetic characteristics as compared to microscopy images taken from vascular plexus in quail embryos. Moreover, we showed that, in this model, cell elongation arises as an emergent consequence rather than a driving force and that pattern formation is possible for realistic values for the VEGF diffusion coefficient.

#### 7.4.2 Experimental validation and quantitative modeling

The described paracrine chemotaxis model poses several predictions that can be tested experimentally. First, network formation is predicted to be abrogated in the absence of paracrine or externally attributed VEGF. Second, external VEGF is predicted to accumulate in the regions surrounding endothelial cells, and to colocalize with extracellular matrix molecules such as fibronectin and heparan sulphates. Third, the dynamics of VEGF in the pericellular regions are predicted to be dominated by binding and unbinding rather than diffusion.

In collaboration with Takashi Miura and colleagues, we carried out experiments to validate these predictions in an *in vitro* assay of HUVECs in Matrigel, presented in chapter 9. The first prediction was tested using a Endothelial Growth Medium (EGM-2) lacking a VEGF supplement, resulting in a severely reduced ability for network formation. Although this result does not exclude a role of VEGF as required for survival or stimulating chemokinesis, it is in line with its predicted effect under the paracrine chemotaxis model. The pericellular localization of external VEGF was confirmed by application of fluorescently labelled VEGF to an established network assay of HUVECs, which allowed us to visualize the spatial distribution of external VEGF using fluorescence microscopy, leaving endogenous VEGF, if any, invisible. Additionally, immunohistochemistry for the two extracellular matrix molecules showed localization of these molecules in the pericellular areas. To investigate the dynamics of VEGF, a fluorescence recovery after photobleaching (FRAP) was used to compare the dynamics in cell-free Matrigel and in the vicinity of cells. Analysis of these recovery curves not only confirmed the reaction-dominated VEGF dynamics near cells, it also provided quantitative estimates for the VEGF diffusivity in Matrigel as well as the binding and unbinding rate of VEGF to matrix molecules. These estimates were combined with data about the molecular degradation obtained by performing an ELISA assay in cell-free conditions.

After validation of the prediction of the paracrine chemotaxis model, we used the quantitative estimates about VEGF kinetics that we obtained from the above experiments to verify whether network formation was possible in our mathematical model for these biophysically realistic parameters. After calibration of the chemotactic sensitivity parameter using published data on HUVECs in a microfluidic device<sup>262</sup>, we were able to show network formation in this quantitative version of the paracrine chemotaxis model.

## 7.5 CONCLUSIONS

Vascular network formation in the embryo is a prime example of biological morphogenesis. The *de novo* establishment of reticular network of blood vessels is crucial to organogenesis. It is well-studied experimentally and can be mimicked *in vitro* cultures allowing easy manipulation. For these reasons, it has been the topic of a range of mathematical modeling studied. These studies have shown that cellular networks can be established by a remarkably large number of different mechanisms. In addition to the models described above, it has been shown that networks can also arise by e.g. preferential attraction to elongated cell structures<sup>263</sup> or even by adhesion between elongated cells<sup>264</sup>.

When multiple, *a priori* equally plausible, models are able to reproduce the phenomenon under investigation, experimental validation of model predictions and the establishment of quantitative computational models, becomes indispensable. In this

sense, the work presented in chapters 8 and 9, provide a classic example of the iteration between model and experiment in multicellular systems biology: First, a new mechanism is proposed based on literature on vasculogenesis *in vivo* and formulated in terms of multi-scale computational model that is shown to be able to produce morphometrically realistic networks *in silico*. The predictions generated by this model are subsequently validated *in vitro* using quantitative methods which, in turn, provide the measurements to be used in the establishment of a quantitative *in silico* model.

Nevertheless, these studies have several shortcomings that are left for future investigation. (1) It is assumed, for instance, that endothelial cells are responsible for the accumulation of available VEGF binding sites on ECM molecules in pericellular regions. However, it is not clear whether this is the result of secretion of matrix components by endothelial cells (e.g. fibronectin expression has been detected in endothelial cells<sup>265</sup>), or their biochemical modification (e.g. through RGD-integrin interaction<sup>266</sup>) or by mechanical traction forces generated by endothelial cells (e.g. as detected by traction force microscopy<sup>267</sup>). (2) Since the paracrine chemotaxis model has only been implemented in the cellular Potts model and no comparison with other cell-based models (e.g. subcellular elements models) has been conducted, it remains unclear to what extent the results depend on the particularities of this model implementation, such as the underlying regular lattice. (3) Although the HUVEC in Matrigel assay is a widely used model system for vasculogenesis and angiogenesis, the relevance of this culture system to vascular development *in vivo* is not clear, in particular with respect to the properties of the extracellular matrix and physiological doses of relevant growth factors. (4) Recent time-lapse imaging of vasculogenesis and angiogenesis have revealed a high level of dynamic cell movements within nascent blood vessels<sup>268–270</sup> that is not captured in any of the described mathematical models of vascular network formation. Determining whether or not the proposed mechanisms are compatible with this new dynamic view on blood vessel formation, remains for future work.

---

## EARLY EMBRYONIC VASCULAR PATTERNING BY MATRIX-MEDIATED PARACRINE SIGNALLING\*

---

### 8.1 INTRODUCTION

During embryonic vasculogenesis, the earliest phase of blood vessel morphogenesis, isolated vascular cell progenitors called angioblasts coalesce and assemble into a reticular pattern<sup>256</sup>. Vasculogenesis is the predominant blood vessel growth mode during early embryonic development, forming a protovascular bed known as the primary vascular plexus. Later, including postnatal and adult stages, this is remodeled by angiogenesis into a complex hierarchical and highly efficient transport system composed of arteries, arterioles, veins, venules and capillaries<sup>256,271</sup>.

The primary vascular plexus is characterized by cells forming a polygon-like pattern. This reticular network structure is ubiquitous among vertebrates which suggest that it holds intrinsic developmental properties likely related to morphogenetic plasticity and that the patterning process is tightly regulated both from a molecular point of view as well as in space and time. Although a large number of endothelium-specific markers and growth factors have been identified as crucial for normal vascular development, the mechanisms underlying the patterning and coalescence of angioblasts remain unclear<sup>271,272</sup>.

In recent years, different hypotheses have been proposed to explain vasculogenesis and formalized into mathematical and computational models; these are reviewed elsewhere<sup>241,243</sup>. Of particular interest here is a number of studies where chemotaxis is considered as a plausible mechanism for *in vitro* vascular aggregation and pattern-

---

\* This chapter is based on the publication: Alvaro Köhn-Luque, Walter de Back, Jörn Starruss, Andrea Mattiotti, Andreas Deutsch, José Maria Pérez-Pomares and Miguel A. Herrero, *PLoS ONE*, 6(9): e24175, 2011. Author contribution: Walter de Back conceived and designed the study together with Alvaro Köhn-Luque, José Maria Pérez-Pomares and Miguel A. Herrero. Model construction, simulation and data analysis were performed by Walter de Back and Alvaro Köhn-Luque. Walter de Back and Alvaro Köhn-Luque wrote the paper.

ing<sup>113,250–252,255</sup>. These studies assume that mature endothelial cells seeded in gels produce a chemoattractant, typically identified as VEGF, that provides these cells the spatial cues driving their migration. This autocrine model may provide insight on the *in vitro* setting described above, but does not fit well with reported data on early embryonic vascular formation.

Chemotactic mechanisms are indeed compatible with biological data on the migration of angioblasts and their coalescence to form early blood vessels, as well as with some theoretical principles on the role of molecular signalling gradients<sup>273</sup>. However, the autocrine regulation mechanism, in which endothelial cells stimulate themselves by both producing and responding to growth factors, does not seem to be fully supported by the biological evidence so far reported in the literature. As a matter of fact, angioblasts are known to express receptors for chemoattractants (VEGFR-2 and CXCR-4,<sup>274,275</sup>), but there is no evidence that, in the embryo, they produce biologically significant amounts of their ligands as well (VEGF and SDF-1, respectively)<sup>256,276,277</sup>. Instead, it is known that most relevant pro-vascular signals, including VEGF, are expressed by the adjacent endoderm<sup>256,257</sup>. A further problem concerns an assumption related to the diffusivity of the signalling molecule. Some autocrine models require a slowly diffusing, quickly inactivating chemoattractant in order to produce stable cellular networks<sup>113</sup>. The assumed rate of diffusion in these mathematical models is typically orders of magnitude lower than that reported for most common VEGF isoforms<sup>113,278</sup>. Thus, both the source of VEGF and its biophysical properties assumed in models of autocrine regulation do not fit the reported data on early vascular development.

In view of these problems, we propose an alternative mechanism for vascular patterning in the embryo. We assume VEGF to be a paracrine signalling agent, in accordance with its reported endodermal origin *in vivo*<sup>256,257</sup>. Yet, paracrine signalling seems at odds with tight regulation of fine-grained network patterns. In the absence of additional regulatory mechanisms, diffusive signals from nearby tissues lack the ability to create precise spatially restricted cues. Interestingly, however, angioblasts are known to produce extra-cellular matrix (ECM) molecules that are able to bind pro-vascular growth factors, including VEGF<sup>261,279–281</sup>. These can immobilize diffusive signalling molecules and thereby provide fine-grained spatial motility cues. For this reason, we assume angioblasts produce ECM molecules with VEGF binding domains.

In this paper, we present a mathematical model based on the assumption that binding of paracrine signals to angioblast produced ECM regulates early vascular patterning in the embryo by creating spatially-restricted guidance cues required for directed cell migration and coalescence (see figure 8.1). In the rest of this section, we provide a concise overview of the key biological evidence on the interaction of VEGF and the extracellular matrix that supports these assumptions. To study whether, and under which conditions, network pattern formation is possible under paracrine regulation we introduce a hybrid cellular Potts / reaction-diffusion model. Simulations show that the model accurately reproduces the morphometrics of early *in vivo* vascular net-



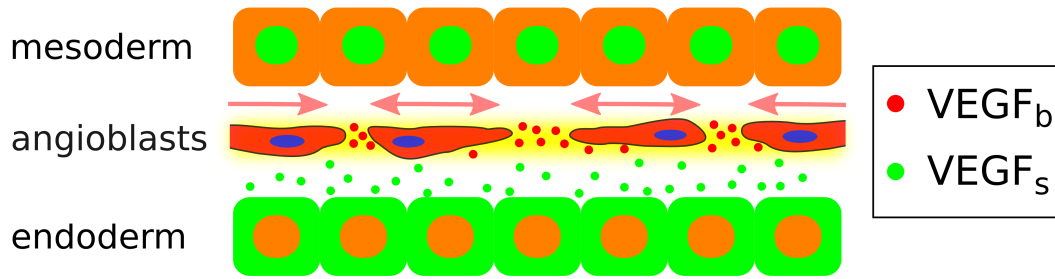


Figure 8.1.: **Paracrine chemotaxis model for vasculogenesis.** Angioblasts (endothelial progenitor cells) are derived from mesodermal cells and assemble into polygonal networks under instructive paracrine signalling provided by the endoderm. Endodermal cells express pro-vascular growth factors such as VEGF. Angioblasts are located in the space between endoderm and mesoderm, surrounded by extracellular matrix (ECM). Angioblasts produce ECM molecules (such as heparan sulphates and fibronectin) with VEGF binding domains (depicted in yellow). This matrix thus acts to store chemotactic growth factors, which provides spatial cues for cell migration.

works in quail embryos. We also check the robustness of that mathematical model by performing a sensitivity analysis with respect to the parameters involved, and explore the dynamics of network formation as well as the role of cell shape and cell density in this process.

### VEGF and the ECM

Angioblasts express a number of endothelium-specific markers like VEGFR-2/FLK-1/KDR; SCL/TAL-1; PECAM/CD31, VE-cadherin or Tie2/Tek<sup>274,282,283</sup>. Some of these molecules are receptors for growth factors essential to vascular morphogenesis, VEGF-A being the most relevant example of a pro-vascular secreted factor considered in the literature. VEGF-A is a glycoprotein with several isoforms arising from the alternative splicing of a unique gene. Splice variants differ in their ability to bind the ECM. Smaller forms are able to diffuse through the ECM, while larger forms bind to heparan sulphate and other characteristic ECM domains with high affinity, thus modulating the distribution of the growth factor in the tissues<sup>261,284–286</sup>.

VEGF signalling is believed to be at the core of the vascular patterning process as suggested by multiple studies<sup>277,287,288</sup>. VEGF expression has been detected in the whole endoderm and some other mesodermal cells, but has not been unambiguously reported in angioblasts<sup>257,289</sup>. As a matter of fact, angioblasts express a variety of VEGF receptors, of which VEGFR-2 is the most relevant one for vascular development<sup>290</sup>. Although different routes can be identified in the transduction of the signal by angioblasts, that one including p38MAPK-MAPKAPK2/3-HSP27 is key in mediating

patterning, as it regulates angioblast protrusive activity towards the signal, thereby modulating cell shape and motility<sup>291</sup>. It has been argued that heparin-binding VEGF isoforms provide spatially restricted cues that polarize and thereby guide sprouting endothelial cells<sup>292</sup>. This suggests a more active role of the ECM in vascular development than previously thought. Indeed several recent reports indicate that the ECM is pivotal to a variety of morphogenetic events including blood vessel formation<sup>125,293</sup>. The specific mechanisms through which the ECM influences angioblast and endothelial cell functions are complex and involve both external structural support and regulation of multiple signalling pathways. Such mechanisms include modulation of growth, differentiation, migration, determination of cell shape and survival, as well as providing storage of growth factors<sup>259–261</sup>.

From the above, it is clear that the ECM can provide crucial cues for early vascular developmental events, but its involvement in such processes tends to be overlooked. This is surprising since it is known that vascular abnormal patterning is characteristic of many animal models where expression/function of a variety of ECM molecules is deficient<sup>294</sup>. A highly relevant finding obtained in knock-out screens is that fibronectin, an ECM glycoprotein essential to the migration of multiple cell types, was found to be the most relevant molecule in early vascular development, closely followed by elements of the VEGF and retinoid signalling pathways and some cell-matrix adhesion molecules<sup>294</sup>. Interestingly, VEGF can bind fibronectin and other ECM molecules and this may affect its downstream effect on cells<sup>227,281,295</sup>. All these findings motivate our study to investigate mechanisms in which the ECM, and matrix-binding VEGF isoforms can lead to coalescence of angioblasts into a polygonal vascular network.

## 8.2 METHODS

### Mathematical model

Before turning to the mathematical formulation of our model, we summarize our assumptions and present the published biological data on which these are based.

(1) A significant number of early embryonic blood vessels develop from isolated angioblasts distributed in the thin, planar surface of the lateral splanchnic mesoderm<sup>256</sup>. Therefore, we can reduce the modeling problem to a 2D domain. We neglect previous developmental events related to the origin of angioblasts, like the specification and commitment of mesodermal cells to the endothelial lineage.

(2) VEGF is known to be produced mostly in the endoderm<sup>257</sup>, and binds to various ECM molecules such as fibronectin and heparan sulphates<sup>261,281</sup>. We model growth factors as diffusible molecules that become non-diffusive when bound to ECM molecules following the mass action law.

(3) Early vascular cells are known to migrate in an ECM rich in fibronectin and heparan sulphates, that can be deposited by endothelial cell precursors themselves<sup>261,279,280</sup>.

Likewise, in our model, we assume angioblasts to produce directly or indirectly non-diffusive ECM components able to bind VEGF.

(4) ECM-bound VEGF provides stronger signalling cues for cell motility or shape remodeling than the freely diffusive forms<sup>227,292</sup>. Consequently, in our model motility is strongly biased towards upward gradients of anchored growth factors rather than to gradients of the freely diffusive forms.

(5) Binding of VEGF to heparan sulphates in ECM molecules offers protection to growth factors against enzymatic degradation<sup>284,296</sup>. Therefore, we neglect degradation or removal of bound growth factors in the model.

(6) Although it is commonly accepted that endothelial cells have a low proliferation rate in the adult and a high one in the embryo, the reported proliferation rates for angioblasts *in vivo* are not high<sup>297–299</sup>. Moreover, administration of exogenous VEGF in the embryo results in dysmorphogenesis by hyperfusion, rather than increased cell numbers<sup>300</sup>. Here we assume proliferation not to be a major limiting factor during the first steps of early vascular formation, and do not include cell proliferation or cell death in our model.

Once particular hypotheses have been established, a mathematical formalism which unambiguously describes them has to be chosen. Previous studies have used different mathematical methods to model vasculogenesis. Most prominently, continuous methods such as differential equations<sup>248,251,263,301</sup>, and discrete approaches with spatially extended cells<sup>113,252,302</sup> have been suggested. While continuous models are appropriate to account for spatio-temporal dynamics of large systems, discrete cell-based methods appear as a useful choice to describe the dynamics of small populations, or to link local microscopic behaviors with macroscopic, collective ones.

In this study, we adopt and modify a hybrid cell-based/continuous model from previous work by Merks and coauthors<sup>252</sup> which represents the multi-scale nature of the problem under consideration and explicitly accounts for cell shape. This model consists of two interconnected modules. On the first one, angioblasts are represented as discrete and geometrically extended objects using a cellular Potts model (CPM)<sup>111</sup>. On the other one, growth factors and ECM molecules are modeled as continuous fields whose distribution is governed by partial differential equations (PDEs).

Let us consider this second module first. Bearing in mind the diffusivity and binding reactions of growth factors and ECM molecules, we propose the following system of differential equations:

$$\begin{aligned}\frac{\partial u}{\partial t} &= D\Delta u + \gamma_1 - f(u, v) - \delta u \\ \frac{\partial v}{\partial t} &= \gamma_2 \delta_{\sigma_{x,0}} - f(u, v) \\ \frac{\partial w}{\partial t} &= f(u, v)\end{aligned}\tag{8.1}$$

where  $u$ ,  $v$  and  $w$  denote the concentrations of soluble VEGF, ECM molecules with free heparan binding domains and bound VEGF, respectively. Parameter  $D$  is the diffusion coefficient of VEGF,  $\gamma_1$  is the constant rate of VEGF production,  $\gamma_2$  is the production rate of ECM molecules by cells ( $\delta_{\sigma_x,0} = 1$  inside cells and  $\delta_{\sigma_x,0} = 0$  outside cells). We assume VEGF-heparan sulfate interaction takes place according to mass action with second order kinetics, so that  $f(u, v) = \alpha uv$  with effective kinetic rate  $\alpha$ . Note that this simple reaction-diffusion model does not include terms for saturated production, enzymatic kinetics or cooperative binding. We thus focus on the feedback mechanisms between the molecular and cellular levels, coupled through ECM-mediated chemotaxis, and reduce to a minimum the assumptions on the underlying biochemical kinetics, about which little is known.

Cells are modeled using a CPM in which each of the  $N$  cells to be tracked is represented by a connected subdomain of a 2D square lattice. The same index  $\sigma = \{1, 2, \dots, N\}$  labels all the lattice sites of a particular cell while the special index  $\sigma = 0$  labels the medium, i.e. all lattice sites not occupied by cells. In this formalism a cell has finite volume and deformable shape. The interfaces between two different lattice sites  $x$  and  $x'$  with different indexes  $\sigma_x \neq \sigma_{x'}$  represent membrane boundaries between cells or between cells and the ECM. To each of these boundaries, a characteristic binding energy is assigned:  $J_{cc}$ , when the interface is between two different cells and  $J_{cm}$ , when it lies between a cell and the surrounding ECM. An energy penalty increasing with the cell's deviation from a selected target area  $A_\sigma$  imposes an area constraint on the cells.

The corresponding Hamiltonian is defined as follows:

$$H = \sum_{\{x, x'\}_n} J_{\tau(\sigma_x)\tau(\sigma_{x'})}(1 - \delta_{\sigma_x, \sigma_{x'}}) + \lambda \sum_{\sigma > 0} (a_\sigma - A_\sigma)^2 \quad (8.2)$$

where  $\tau(\sigma)$  represents the type of object occupying a grid space  $\sigma$ , which in this case can only be angioblast (c) or medium (m). The term  $(1 - \delta_{\sigma_x, \sigma_{x'}})$  ensures binding energies are only considered between different cells. The term  $(a_\sigma - A_\sigma)$  represents a cell's deviation from its target area and  $\lambda$  represents the cell's resistance to such deformations. The first summation is taken for the  $n^{\text{th}}$ -order neighbors in each lattice site. The second summation goes for all cells with the exception of the medium.

Cell dynamics are generated in the CPM by a modified Metropolis algorithm. The latter randomly chooses a lattice site,  $x_{\text{target}}$ , and computes what the difference in energy,  $\Delta H$ , would be if a randomly selected neighboring site,  $x_{\text{source}}$ , would copy its state into this site. The probability of accepting the change,  $P$ , depends on the difference in the energy costs:

$$P(\Delta H) = \begin{cases} 1 & \text{if } \Delta H < 0 \\ e^{\frac{-\Delta H}{T}} & \text{otherwise,} \end{cases} \quad (8.3)$$

so that cell extensions that diminish  $H$  are given priority. In this way, the cell shape is updated locally. Parameter  $T$ , henceforth referred to as fluctuation energy, is a biological analogue to the energy of thermal fluctuations in statistical physics<sup>303</sup> and it is considered here as a measure of cell motility. Chemotaxis is modeled as a bias in the direction of higher VEGF concentrations. More precisely, in our case  $P$  is considered to depend on:

$$\Delta H_{\text{chemotaxis}} = \Delta H - \mu_b(w(x_{\text{target}}) - w(x_{\text{source}})) - \mu_s(u(x_{\text{target}}) - u(x_{\text{source}})), \quad (8.4)$$

where a distinction is made between bound and soluble VEGF. The strength of signalling provided by the bound and soluble forms can be varied by setting the parameter  $\mu_b$  and  $\mu_s$  while preserving their total value  $\mu_t = \mu_s + \mu_b$ . Stronger chemotaxis towards bound VEGF is accounted for by setting  $\mu_b > \mu_s$ . Finally, the unit of time in the simulation is defined as one Monte Carlo step (MCS). One MCS corresponds to the number of random update attempts equal to the number of lattice sites.

### Experimental images

Blood vessels of quail embryos are labeled with the QH1 antibody (endothelial membrane) using DAPI for nuclear counterstain, at 36-40 hours after incubation. Confocal microscopy allows us to clearly visualize the developing vasculature and all the nuclei in the tissue. Images are then taken of wide areas of the embryo where angioblasts have just assembled into networks (see figure 8.2). QH1 labeling of the cellular structures is achieved by segmentation of the appropriate fluorescence channel. This results in a binary image where vascular zones are labelled with one and avascular zones or lacunae are labelled with zero. Segmentation of DAPI fluorescent stain, marking all nuclei in the confocal plane, is used to estimate the angioblast cell density by co-localization of identified nuclei with the QH1 membrane label using the segmented binary image.

Our simulation software produces binary images in which cells and lacunae are labelled as described before so that both series can be examined and compared by means of the same morphometric techniques.

### Morphometric methods

Experimental and simulated networks represented in binary images are characterized by quantifying the lacunae (number, size and roundness of avascular zones), the vascular structures (number of connected components, coverage and widths), network properties (number of nodes, degree distribution, percolation and spanning length) and fractal properties (fractal dimension and lacunarity). Image processing and statistical

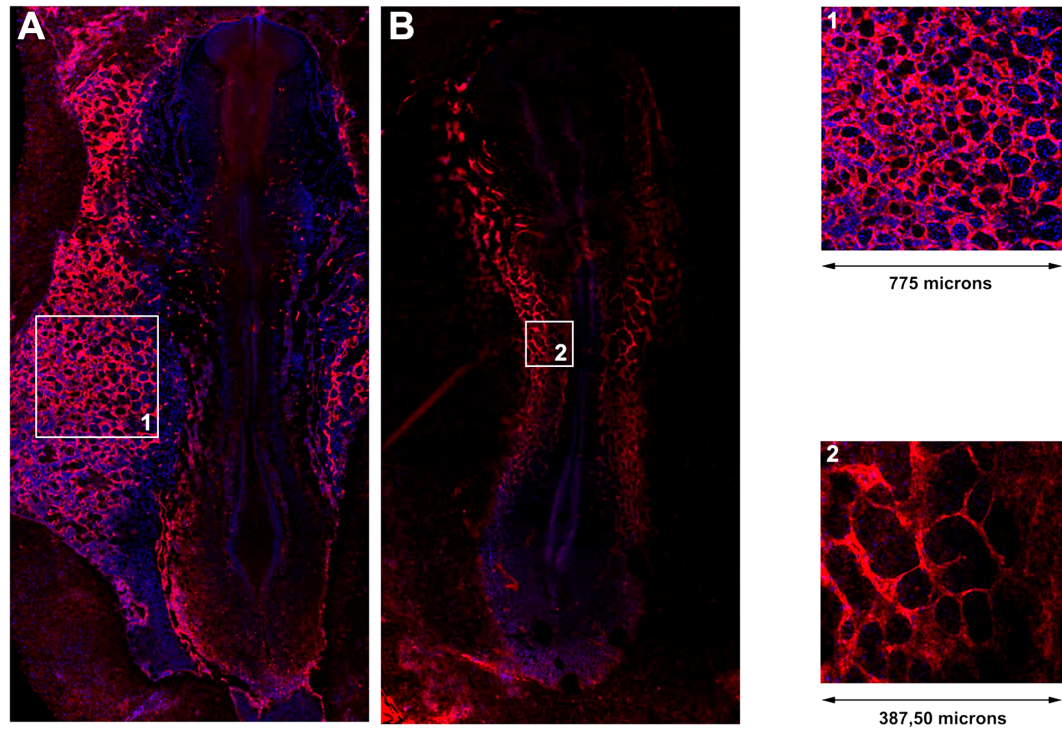


Figure 8.2.: *In vivo* vascular network in the quail embryo. Laser confocal microscope reconstruction of the extra-embryonic (A) and intra-embryonic (B) vasculature of early quail embryos (36-40 hours of incubation). Embryonic blood vessels are identified by the QH1 antibody (red). Cell nuclei have been counterstained with DAPI (blue). The inserts depict extra-embryonic (1) and intra-embryonic (2) vascular networks in more detail; the former are used in analysis and validation of the mathematical model.

analysis have been performed in Matlab 2009b using morphological functions supplied by the Image Processing Toolbox (version 6.4) supplemented by custom-made routines. Together, the set of morphometric features provides us with detailed quantification of vascular patterns obtained *in vivo* and *in silico*.

The number of connected vascular zones and lacunae are two key network characteristics and correspond to the 0<sup>th</sup> and 1<sup>st</sup> Betti number used in algebraic topology (see<sup>304</sup> and references therein). These are identified by tracing the exterior boundaries of objects in the binary image. The sizes of lacunae and vascular structures equal the number of pixels they occupy, while the roundness of lacunae is measured using the isoperimetric quotient, defined as the ratio of the measured area to that of a circle having the same perimeter,  $R = 4\pi A_l / P_l^2$ , where  $A_l$  and  $P_l$  are the lacuna area and perimeter, respectively. The coverage of angioblasts is recorded as the ratio of the number of nonzero pixels to the size of the binary image. The network is said to be percolative when a vascular structure exists that spans the image both horizontally and vertically.

Assessing some of the morphometric features requires computation of a skeletonization of the network, using a morphological operation known as thinning. Thinning reduces the network structures to a skeletal remnant that largely preserves the extent and connectivity of the original one while pruning away redundant foreground pixels. The spanning length of the network is then measured as the sum of nonzero elements. The widths of cellular cords are assessed by measuring the Euclidean distance from each nonzero pixel on the thinned structure to its closest lacuna. Nodes in the thinned structure are usually identified as nonzero points with 3 or more nonzero pixels among the 8 adjacent sites, indicating a point at which the structure branches. Although this method is often used<sup>252,305</sup>, it actually overestimates the number of nodes when no further processing is applied. Therefore, we corrected the measurement by merging proximate nodes whenever the distance between them is smaller than the distance from the node to the nearest lacuna. This procedure reduces the number of nodes on the thick vascular segments while preserving them on thin cords. Moreover, using this method we gain information on the degree distribution, i.e. on the distribution of connections or edges a node has to other nodes. Finally, we performed fractal analysis to quantify the complexity and space filling properties of network patterns. Informally, the fractal dimension  $D_f$  measures how a fractal-like structure fills the space for decreasing scales, while the lacunarity  $L_f$  assesses its 'gappiness', i.e. the distribution and size of the empty domains<sup>306</sup>. This has previously been applied to characterize vascular growth in e.g. chick chorioallantoic membrane (CAM) angiogenesis<sup>305,307</sup>. The fractal properties of the the thinned network structure were estimated by applying the box counting methods and gliding box method<sup>308,309</sup>, using the FracLac plugin for ImageJ<sup>310</sup>.

Parameter	Symbol	Value	Unit	Source
<i>PDE</i>				
Diffusion coefficient VEGF <sub>s</sub>	D	10	$\mu\text{m}^2/\text{s}$	<sup>278</sup>
Production rate VEGF <sub>s</sub>	$\gamma_1$	$10^{-3}$	a.u./ $\mu\text{m}^2/\text{s}$	est.
Production rate ECM	$\gamma_2$	$10^{-3}$	a.u./ $\mu\text{m}^2/\text{s}$	est.
Binding rate VEGF <sub>s</sub> +ECM	$\alpha$	$10^{-1}$	a.u./ $\mu\text{m}^2/\text{s}$	est.
Degradation rate VEGF <sub>s</sub>	$\delta$	$10^{-2}$	$\text{s}^{-1}$	est.
<i>CPM</i>				
Fluctuation energy	T	50		<sup>252,254,255</sup>
Target area	$A_\sigma$	95	pixels	emp.
Cell rigidity	$\lambda$	25		<sup>113,252</sup>
Cell-cell binding energy	$J_{cc}$	20		<sup>252</sup>
Cell-medium binding energy	$J_{cm}$	10		<sup>252</sup>
Chemotaxis strength	$\mu_t = \mu_s + \mu_b$	2000		<sup>252,255</sup>
Signalling strength VEGF <sub>b</sub>	$\mu_b/\mu_t$	0.75		<sup>227,292</sup>
Signalling strength VEGF <sub>s</sub>	$\mu_s/\mu_t$	0.25		<sup>227,292</sup>

Table 8.1.: **Model parameters for equations 8.1, 8.2, 8.3, 8.4.** “est.” refers to estimated parameters, “emp.” means empirically measured.

### Simulation set-up

When selecting simulation scenarios, several choices are determined by the experimental setup. For instance, the size of the lattice where CPM and PDE are simulated is adjusted to resemble  $775 \times 775 \text{ mm}^2$  experimental images. Specifically, we consider a square lattice of  $400 \times 400$  pixels, where each lattice node represents  $2 \text{ mm}^2$ , thus giving a total simulated tissue of  $800 \times 800 \text{ mm}^2$  or  $0.64 \text{ mm}^2$ . Cell densities are estimated from experimental images by co-localization of nuclei with the vascular structure, resulting in an average density of  $\approx 1750 \text{ cell mm}^2$  (results not shown), which is used as a reference in the simulations ( $1100 \text{ cells over } 800 \times 800 \text{ mm}^2$ ). As initial condition, cells are distributed over the lattice in either a regular or a semi-random mesh. The latter is constructed by randomly displacing cells from the regular mesh, where displacements are smaller than one cell diameter.

Parameter choices in the CPM module have been made as in previous works<sup>113,252,254</sup>, see table 8.1. In particular, we neglect surface tension between cells and the ECM for sake of simplicity, by setting  $\gamma_{cm} = J_{cm} - J_{cc}/2 = 0$ .

The criteria for selecting PDE module parameters are different. A value for the diffusion coefficient of VEGF has been taken similar to those reported in the literature<sup>278</sup>,



although we shall see later that our results are largely insensitive with respect to this coefficient. The rest of the parameters has been chosen on the basis of estimated time scales of processes. For instance, binding is fast relative to VEGF and ECM production, and the corresponding parameters are selected so as to fit experimental data (see table 8.1). Then, a sensitivity analysis on these parameters is performed to give us an idea of the expected variability in terms of some of the quantified morphometric properties.

Each MCS in simulation corresponds to one second. For this choice, the total simulated time is about one hour, agreeing well with the estimated time scale in which the process takes place *in vivo* (from minutes to a few hours).

Simulations have been implemented using our own C++ based modeling environment Morpheus. Multiple simulation repetitions are performed ( $n = 10$ , unless stated otherwise) with different random seeds, of which the mean and standard deviation are shown. Figure 3.8 lists the full model description of the simulation model in MorpheusML. This model is also available in the Supplementary Online Material under <http://walter.deback.net/thesis>.

## 8.3 RESULTS

### 8.3.1 Model simulations yield vascular-like reticular patterns

We compared vascular networks obtained from quail embryos *in vivo* to those resulting from model simulations along a series of morphometric properties, summarized in figure 8.3. Examples of binary images used in this comparison are shown with their thinned skeleton (red pixels) and detected nodes (blue circles) in panel a. We found that the morphometrics of simulated networks correspond well to experimental observations, as shown in the statistical comparison in figure 8.3B. Despite small deviations, the number of nodes and edges, network length and interface length, as well as fractal properties and coverage measured in the simulated networks are remarkably similar to the corresponding values obtained from experimental images. Lacunarity, a measure for lack of translation and rotational invariance, is observed to be slightly higher in the *in vivo* situation. This may be due to the spatial heterogeneity in tissue density along the mediolateral axis, which is not reflected in our model. Although the number of cells used in simulation has been estimated from experimental images, the observed coverage (the relative size of the area covered by angioblasts) is also seen to be slightly higher *in vivo*. The differences in network and interface length are related to the previous feature, since denser vascular structures have more and longer edges at the cost of cell-lacuna interface length.

In order to further characterize the vascular network, we quantified distributions of several measured morphometric properties, shown in figure 8.3C. As a result of our procedure to merge proximate nodes, we recovered the number of connections per node, known as the node degree (indicated as size of blue circles in figure 8.3A). Al-

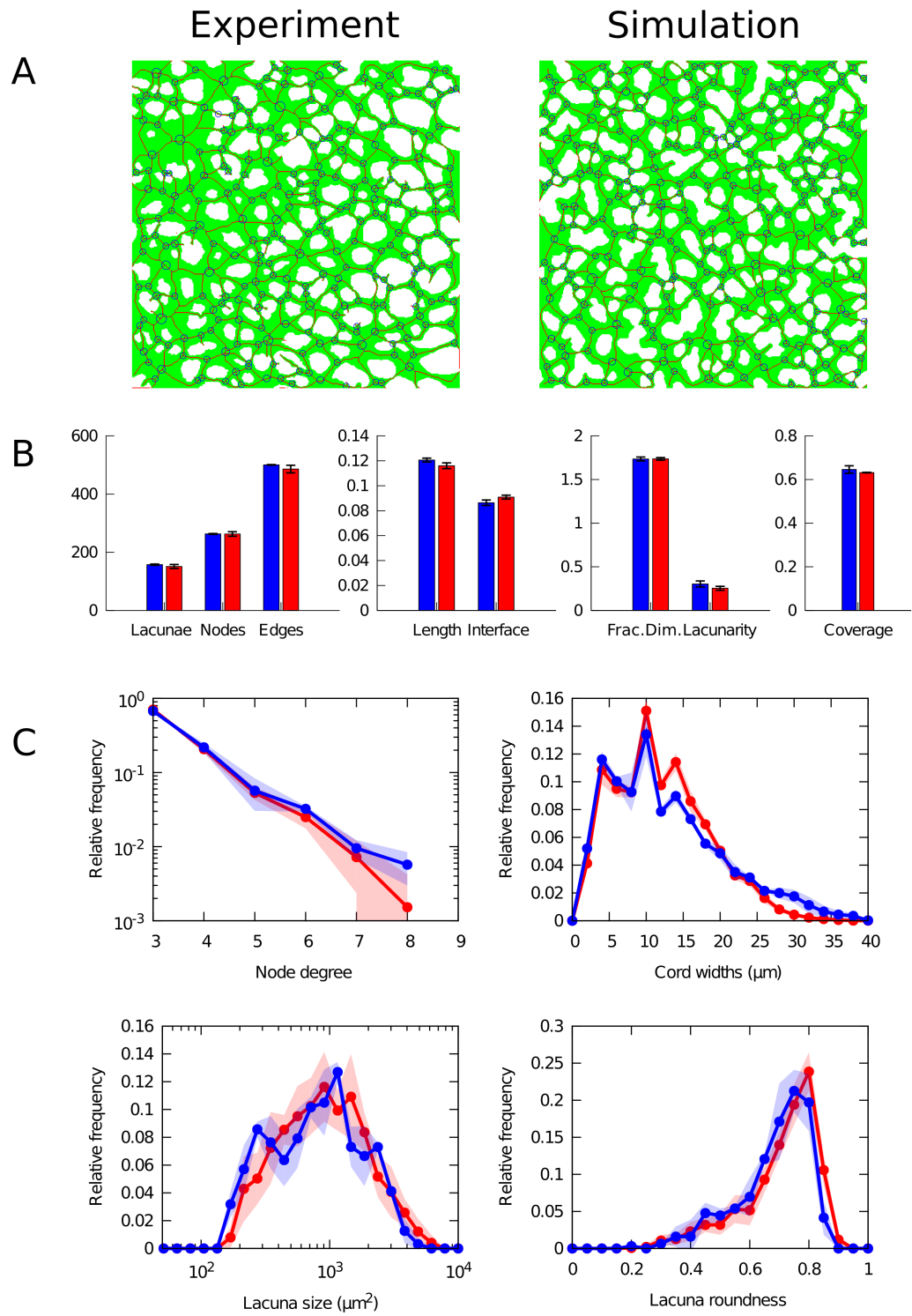


Figure 8.3.: **Morphometric comparison.** (Caption on next page.)

Figure 8.4.: Comparison between experimental (in blue) and simulated (in red) vascular networks (after 3000 MCS). (A) Binary images over cellular structures (green) overlaid with skeletonized network (red), detected branching points (blue points) and corrected nodes (blue circles). (B) Morphometric statistics. Boxes show average values ( $n=2$  for experiments;  $n=10$  for simulation) and error bars indicate standard deviation. (C) Distributions of morphometric properties. Lines show average values; filled areas indicate standard deviations.

though the range of observed node degrees is rather small, an exponential decay can be recognized in their distribution, with high frequency of small degree and rare high-degree nodes. Using both the skeletonized structure and the underlying binary image, we assessed the distribution of cord widths by measuring the shortest distance of each pixel on the skeletonized image to an avascular region. The distribution thus obtained reveals a right-skewed distribution of cellular cord widths with thin cords forming the highest frequency, while thick cords are relatively rare. Simulated networks display less right-skewness, which may again be related to the mediolateral heterogeneity in the experimental images. Nevertheless, the modes of both distributions coincide and both show a distinct pattern of high and low frequencies around the mode. These peaks are caused by the discreteness of the number of cells involved in a cord, a fact that is captured by the discrete cell-based submodel. Interestingly, the mode of the distribution ( $10\mu\text{m}$ ) is well below the diameter expected for isotropic cells which indicates a strong contribution of elongated cells (discussed in more detail below). Finally, we compared the size and shape distribution of lacunae. In both experiment and simulation, a wide lognormal distribution can be discerned in lacuna size distribution, spanning almost two orders of magnitude around the most frequent size at approx.  $1000\mu\text{m}^2$ . The distribution of roundness of lacunae is left-skewed with most holes being nearly circular. Note that these two quantities are related to each other by the fact that larger holes tend to have a less circular shape. Lacunae size and roundness are thus negatively correlated, although this effect is more clearly seen in simulations (Pearson correlation  $r = -0.81$ ) than in the segmented experimental image ( $r = -0.27$ ).

In summary, our detailed morphometric analysis shows that a model based on chemotaxis towards growth factors bound to ECM secreted by angioblasts is able to produce reticular patterns as those observed in vascular networks in quail embryos. We chose to describe and compare the vascular networks using a broad spectrum of morphometric properties for two reasons. First, detailed quantitative morphological characterization of early vascular networks is needed to characterize normal vascular development. It can also prove helpful to detect defects in experiments in which specific processes are perturbed. On the other hand, a fact relevant for our study here is that, in order to ascertain geometrical and functional similarities, multiple shape descriptors are required to compare vascular shapes. In principle, this represents a difficulty since an accurate fit with respect to one shape parameter (e.g. number of nodes)

needs not be compatible with a similar fit with respect to another such parameter (e.g. distribution of cord widths). A remarkable fact that results from our simulations below is that a good fitting can be obtained simultaneously with respect to a large number of morphometric parameters.

### 8.3.2 Sensitivity analysis

In order to check the robustness of the model, we investigated the sensitivity of the simulation results with respect to their parameters. First, a sensitivity analysis was performed on the coefficients specifying the rates of production ( $\gamma_1$  and  $\gamma_2$ ), binding ( $\alpha$ ) and degradation ( $\delta$ ) of VEGF and ECM, respectively. Simulations were performed in which these parameters are varied independently over two orders of magnitude from their reference values (see table 8.1). We measured the change in morphometric properties with respect to the reference simulation described above, shown in figure 8.5A. We find that a 10-fold decrease and increase of the VEGF-ECM binding rate  $\alpha$  and VEGF decay rate has only moderate or negligible effects on the network morphology. However, variation of the VEGF and ECM production rates ( $\gamma_1$  and  $\gamma_2$ ) does have large impact on the structure of the network. Upon decrease of these production rates a marked decline in the number of lacunae is observed, indicative of vascular dysmorphogenesis. This is readily understood under our model assumptions, since effective removal of either VEGF or ECM disrupts binding of growth factor to the matrix, which normally provides the spatial cues for cellular patterning. The analysis further shows that a 10-fold increase in VEGF production ( $\gamma_1$ ) has little impact on network formation. Interestingly, however, a similar increase in matrix production ( $\gamma_2$ ) does disturb proper morphogenesis. This happens because overproduction of ECM allows fast accumulation of the bound VEGF. Due to chemotaxis towards bound VEGF, this causes premature immobilization of cells and hampers the coalescence of cells and their assembly into a polygonal pattern.

We further explored the sensitivity of the results with respect to VEGF diffusivity. Previous model of vasculogenesis were dependent on unrealistically slow diffusion of VEGF, for instance in<sup>252</sup>. To understand the role of VEGF diffusion in our model, we performed simulations with different scenarios concerning VEGF diffusion: no diffusion ( $D = 0$ ), normal diffusion ( $D = 10 \mu\text{m}^2/\text{s}$ ) and a well-mixed system ( $D = \infty$ ). The latter is modeled by redistributing the total amount of VEGF homogeneously over the lattice after every reaction step. Remarkably, the results in figure 8.5B show that our model is extremely robust against variation in that parameter. We explain this virtual independence on VEGF diffusivity by reference to the assumed homogeneous production of growth factor over the modeled area. Therefore, even in the absence of diffusion, binding of VEGF to ECM will occur throughout the tissue and guide motility. On the other hand, under high diffusivity, VEGF is not only produced but also

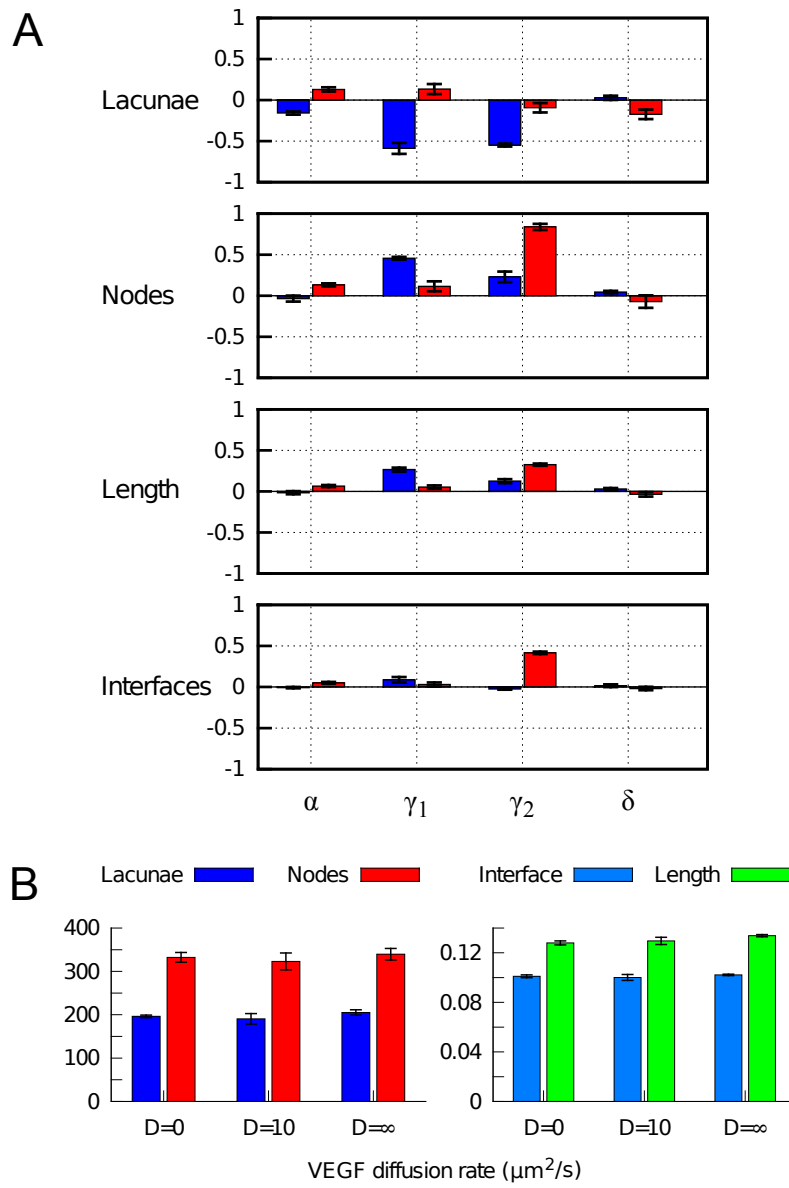


Figure 8.5.: **Sensitivity analysis of PDE parameters.** (A) Sensitivity to rates of binding ( $\alpha$ ), VEGF and ECM production ( $\gamma_1$  and  $\gamma_2$ ), and degradation of soluble VEGF ( $\delta$ ). Changes in various morphometric properties were measured for simulations ( $n=3$ ) in which each parameter was independently varied by a 10-fold decrease (blue) and a 10-fold increase (red). (B) Sensitivity to VEGF diffusivity. Morphometric quantities are shown for simulations ( $n=3$ ) with non-diffusive VEGF ( $D = 0$ ), with normal VEGF diffusion ( $D = 10$ ), and with well-mixed VEGF ( $D = \infty$ ).

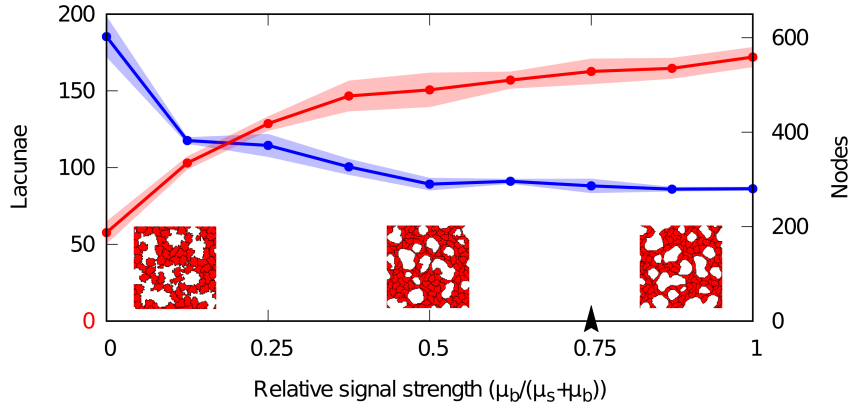


Figure 8.6.: **Sensitivity analysis of chemotactic signal strength.** Sensitivity of morphometric parameters to relative strength of bound ( $\mu_b$ ) and soluble ( $\mu_s$ ) VEGF. Red points (lacunae) and blue points (nodes) show averages of measured quantities in simulations ( $n = 3$ ), half-transparent regions represent standard deviations. Insets show portions of networks ( $200 \times 200 \mu\text{m}^2$ ) where the relative signalling strength ( $\mu_b/\mu_t$ ) is set to soluble-VEGF-only ( $\mu_b = 0$ , left), bound-VEGF-only ( $\mu_s = \mu_t$ , right) and equal strengths ( $\mu_b = \mu_s$ , center). Arrowhead indicates the reference value.

redistributed in instantaneously homogeneous manner, leaving matrix binding and motility relatively unaffected.

Additionally, we investigated the dependency of the resulting network pattern on chemotaxis. Based on experimental evidence<sup>227,292</sup>, we assume that the ECM-bound VEGF provides a stronger signalling cue than soluble VEGF. Accordingly, cells in our model have a stronger chemotactic response to bound VEGF ( $\mu_b$ ) than to soluble VEGF ( $\mu_s$ ) in a 3 : 1 ratio. To ensure that network pattern formation is not dependent on our particular choice of these parameters, we varied the relative strengths of chemotaxis towards bound and soluble VEGF. Effectively, this alters the relative strength of the signalling cues on cells. Figure 8.6 shows that network formation is perturbed when soluble VEGF is the major signalling agent ( $\mu_s > \mu_b$ ), which is due to a lack of spatially restricted cues. On the other hand, when cells react more strongly to ECM-bound VEGF ( $\mu_b \geq \mu_s$ ), network formation is relatively stable in a broad domain. Our reference values of these parameters (arrowhead in figure 8.6; table 8.1) are within this region.

### 8.3.3 Model dynamics display fast coalescence and slow remodelling

Experimental images are obtained by fluorescent staining which requires fixation of the embryo. Therefore, no temporally resolved data are available that would enable us to track the development of a vascular bed *in vivo*. Using the simulation model,

however, it is possible to look into the dynamics of vasculogenetic development. Although proliferation is known to occur during the assembly of the vascular bed, this is omitted from the model for simplicity. Instead, we assumed a constant population size of 800 cells ( $1250 \text{ cells/mm}^2$ ), a figure that lies between a rough estimate of newly differentiated angioblasts and the cell density measured in our experimental images. The influence of cell density on the network morphology is explored below.

Although the initial population of cells is distributed in a regular mesh, a network pattern forms rapidly during simulation, as shown in figure 8.7. The top panel (figure 8.7A) displays a typical simulation showing the shapes and positions of simulated angioblasts, and the relative concentrations of matrix-bound VEGF. The evolution of the number of connected cellular structures and that of the number of holes in figure 8.7B shows that these quantities move towards a steady state, indicating the dynamic stability of the reticular structure. This stabilization property is in contrast with some models based on autocrine chemotaxis hypothesis<sup>311</sup>.

Further, the dynamics considered in figure 8.7B seem to operate on separate timescales. On the one hand, cells rapidly coalesce from a mostly isolated configuration to form a percolative and closed network consisting of a single connected component (between 0 and 1000 MCS). On the other hand, the number of lacunae changes over a longer time scale, and keeps increasing after all cells have assembled into a single network (between 1000 and 3000 MCS). In other words, subsequently to the assembly of a closed network, cells continue to rearrange and remodel the network and thereby increase the number of lacunae. Substantial remodeling is performed by cells creating new connections across lacunae, resulting in cords with a single cell width (indicated by arrows). Figure 8.7C depicts a close-up of remodeling occurring in the period (1000-2000 MCS) after a percolative network has formed. In addition to increasing the number of lacunae in this way, the resulting holes are smaller and have a more regular and round shape. This remodeling is driven by cell motility over tracks of previously secreted ECM that are subsequently primed by VEGF binding.

#### 8.3.4 Cell elongation results from chemotaxis towards matrix-bound VEGF

Using the same data, we tracked the shape of cells during simulation and quantified their lengths by computing the inertia tensor of an elliptic shape approximation<sup>252,312</sup>. Taking the length of isotropic cells as a reference, we compared the distribution of cell shapes before and after tissue remodeling, as shown in figure 8.9. Before remodeling, cell lengths are skewed towards isotropic length, indicating roundish cells. Afterwards, the position of the mode of the distribution (i.e. the most frequent cell length) has remained unchanged. However, the tail of the distribution has grown, showing that the contribution of elongated cells has markedly increased. Some cells extend up to almost three times the default isotropic length. Importantly, this is not caused by constraining cell shape in our model (as in was done in<sup>252</sup>). Instead, the observed cell elongation

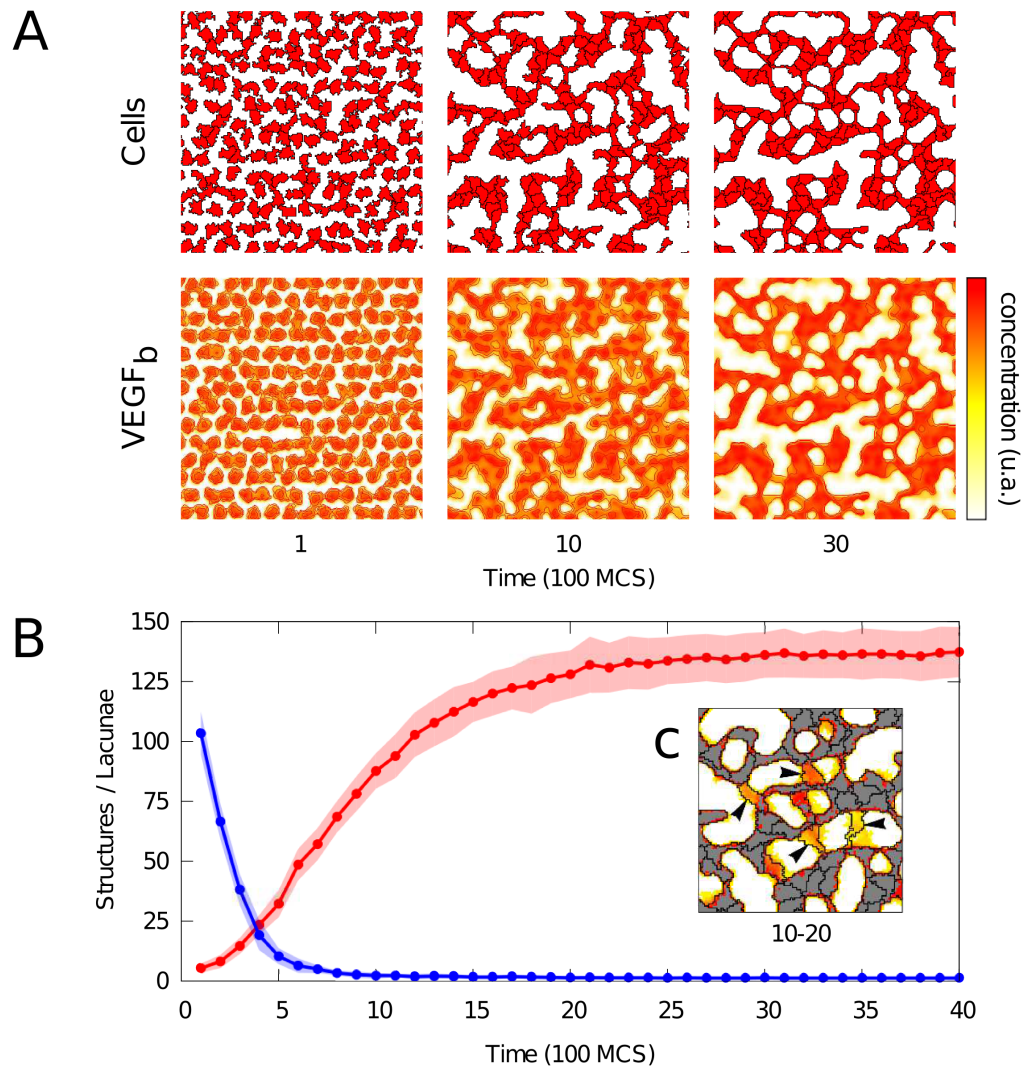


Figure 8.7.: **Dynamics of vascular network formation.**  
(Caption on next page.)



Figure 8.8.: (A) Detail of simulated tissue at various time points, showing the cells (top) and relative concentrations of bound VEGF with isolines (bottom). (B) Dynamics of number of isolated cellular structures (blue) and number of lacunae (red), half-transparent regions indicates standard deviation ( $n=10$ ). (C) Inset depicts remodeling in a small region. This occupancy map is constructed by averaging over binary images in the interval between 1000 and 2000 MCS; lines show cell boundaries at 2000 MCS. Grey/white pixels are cells/lacuna which remained unchanged over this period; colored pixels indicates how long a pixel has been occupied by cells. It shows the creation of new connections (arrows) increasing the number of lacunae. The simulation model is available in the Supplementary Online Material under <http://walter.deback.net/thesis>.

follows as a natural consequence of chemotaxis towards matrix-bound VEGF. Thus, in our model elongation is an effect rather than a cause for the formation of a reticular pattern in the vascular bed.

### 8.3.5 Morphological dependence on cell density

To explore the influence of cell density on the morphology of the vascular network, we performed a set of simulations in which the number of cells is varied, while keeping the simulated tissue size constant, and analyzed the morphometric properties. The results, shown in figure 6.4, point out the presence of various thresholds or maxima at increasing cell densities.

The first threshold appears at a density of  $\approx 900$  cells/mm<sup>2</sup>, where the network exhibits percolation. This percolative threshold is the critical density above which a network first exhibits long-range connectivity. Related to this, the point at which the vascular structure forms a single connected component (so that all cells lie within a single network structure) is found at only slightly higher densities. Both aspects are important for blood transport through the tissue and are established at low cell densities.

At medium cell densities ( $\approx 1500$  cells/mm<sup>2</sup>), we observed a maximum in the number of nodes and the length of cell-lacunae interfaces. A high number of nodes indicates a complex branched network structure. Unsurprisingly, such networks show a maximum in the interface length between vascular and avascular regions. Biologically, it is this surface area between endothelium and surrounding tissue which is crucial to the supply of oxygen and nutrients to that tissue.

When density is further increased ( $\approx 2000$  cells/mm<sup>2</sup>), interface length decreases as the size of the avascular region decreases. Instead, we observe that the number of lacunae and total network length are maximized. This cell density corresponds to that measured in *in vivo* vascular networks in quail embryos at 36-40 hours after incubation.

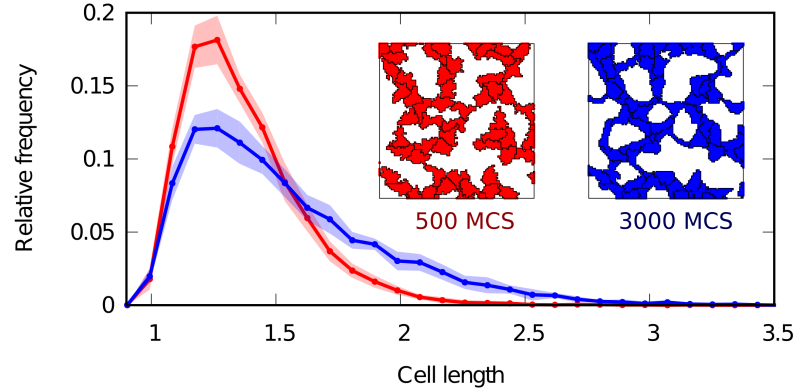


Figure 8.9.: **Cell elongation.** Distribution of cell lengths at different time points during the simulation. Red line and inset depict early, blue depicts late in development. Lengths are normalized to isotropic cells given the target area ( $\sqrt{s4\Lambda_\sigma/\pi} \approx 22\mu\text{m}$ , where  $s = 4\mu\text{m}^2$  is the scaling factor per pixel). Cells become increasingly anisotropic and elongated during vascular patterning. Filled regions represent standard deviation.

#### 8.4 DISCUSSION

We have presented a new mathematical model to describe vascular patterning during early stages of embryonic vasculogenesis. The prevalent view that the formation of vascular network patterns is regulated by autocrine chemotaxis has been formulated largely based on *in vitro* studies. In developing our model, we have instead made use of the current biological knowledge on early vasculogenesis in the embryo. In accordance to the hypotheses made by other authors<sup>256,276</sup>, we have explored the role of paracrine signalling in embryonic vascular development. In particular we have mathematically simulated the role of the extracellular matrix (ECM) in providing spatial cues for angioblasts by storing and/or activating chemoattractive growth factors. Actually, the significance of the ECM during vasculogenesis has long been known<sup>279</sup> and continues to be an active research field both in vasculogenesis and angiogenesis<sup>125,227,261,281,292</sup>, although its capacity for pattern formation by providing signalling cues has not received much attention (see however<sup>227,292</sup>).

Specifically, we combined a 2D discrete cellular Potts model (eq. 9.2 and 8.3) with a continuous reaction-diffusion model (eq. 9.1) under the assumptions that angioblasts are chemotactically attracted towards paracrine VEGF that binds to angioblast-produced ECM. As demonstrated by morphometric analysis, this model is able to produce polygonal cellular patterns that accurately resemble the *in vivo* early vascular bed in quail embryos, recorded by confocal microscopy. The simulated networks show high degree

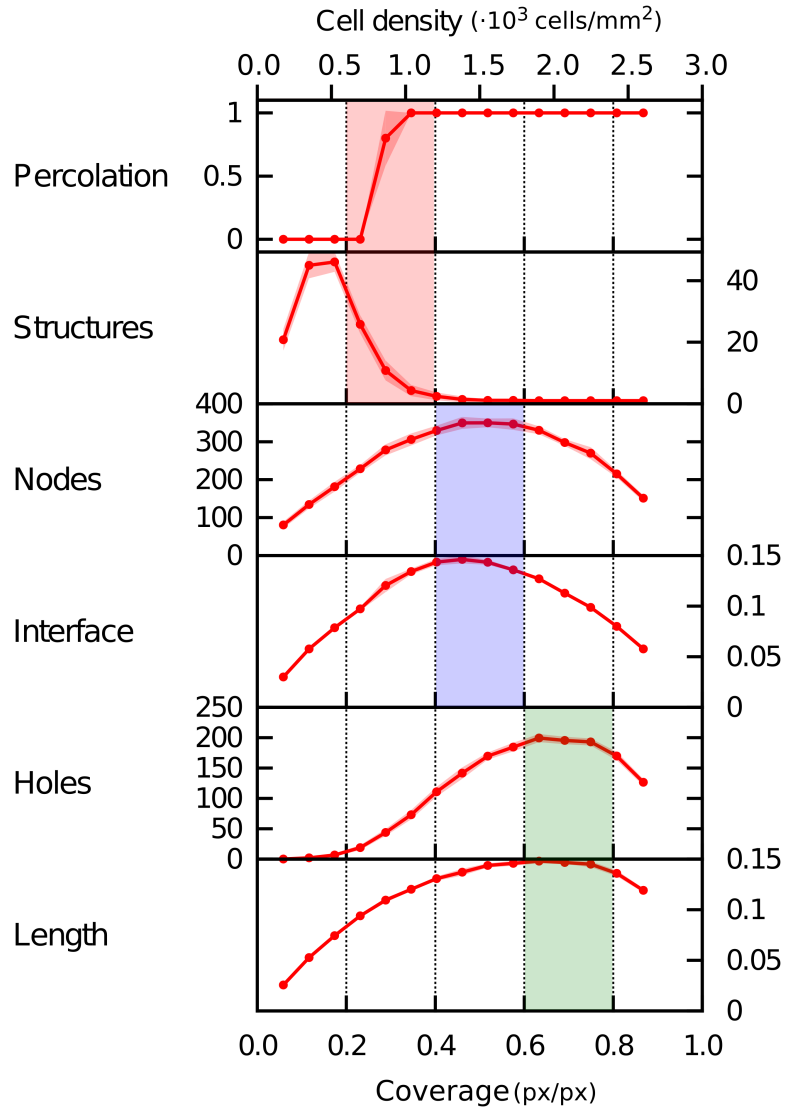


Figure 8.10.: **Morphometric dependence on cell density.** A number of morphometric properties of simulated networks are presented as a function of both cell density (number of cells per area) and coverage (the ratio of angioblasts-covered pixels to the total number of pixels). Averaged over 10 simulation runs, transparency indicates standard deviation. Three optima are shown from top to bottom, at increasing cell densities. Parameters as in table 8.1.

of similarity with respect to a broad spectrum of morphological descriptors, including lacunae number/sizes/shapes, network and interface lengths, cord widths, degree distribution and fractal properties.

At the same time, our model circumvents certain drawbacks of autocrine mathematical models in the context of early embryonic vascular patterning. Indeed, a key problem of the autocrine regulation hypothesis is that no experimental evidence exists that angioblasts produce significant amounts of VEGF, while it is known that the adjacent endodermal tissue produces many pro-vascular growth factors, including VEGF<sup>256,257</sup>. Thus, for the autocrine model to account for early embryonic vasculogenesis, it remains to be explained how embryonic angioblasts could respond more efficiently to low mesodermal VEGF levels rather than to the high amounts of the same growth factor produced by the adjacent endoderm. The mathematical model proposed here does not consider VEGF production by angioblasts, and assumes instead an external source of VEGF. Angioblast-secreted or modified ECM molecules bind and immobilize the paracrine signalling agent in close proximity of cells. Thus, fine-grained spatial cues for chemotactic cell migration can be generated without postulating unrealistically low VEGF diffusion rates<sup>113,252</sup>. Furthermore, the stability of the network structures increases over time, instead of collapsing after a transient time, as in previous models<sup>113,251,311</sup>. Another interesting observation is that cell elongation, a fact which is experimentally observed, needs not be postulated a priori as in<sup>252</sup>. Instead, cells elongate in our model as a natural consequence of chemotaxis towards matrix-bound VEGF.

Simulations of our paracrine model yields several results that are coherent with most of the biological data on embryonic vasculogenesis published up to date. In general, downregulation of VEGF or the ECM molecules where VEGF can bind, as fibronectin or heparan sulphates, is known to severely impair early vascular patterning<sup>261,286,313</sup>. In particular, administration of exogenous soluble VEGF receptors during vasculogenesis, which decreases the level of endogenous VEGF, results in a lack of network formation<sup>287</sup>. The relevance of VEGF matrix binding in providing precise spatial cues has also been addressed in several studies. For instance, embryos that expressed only VEGF isoforms lacking ECM interaction domains lead to greatly reduced vessel branching and network complexity<sup>292</sup>. However, to unravel the interaction between ECM molecules with heparin domains and VEGF matrix-binding isoforms as well as their influence in angioblasts responses, further experimental work is needed.

On the other hand, simulations of the mathematical model here considered, reveals some interesting features of networks formation, which however remain to be experimentally validated. For instance, for a fixed number of cells, pattern formation dynamics on two time scales are observed. First, a fully connected network is formed, so that percolation is attained and middle-long range connectivity is achieved. This can be interpreted as a basic requirement for any functional blood vessel network. On a longer timescale, tissue remodeling takes place, whereby the number of lacunae continues to increase so that the surface area keeps increasing too, thus ensuring an efficient

distribution of nutrients and waste removal. On the other hand, when cell densities are varied, three maxima in the morphometric quantities can be distinguished. At increasing cell densities, these mark the onset of percolation, the attainment of optimal interface length, and that of a maximal number of lacunae, a fact related to the total spanning length. Interestingly, the densities measured in *in vivo* vascular networks correspond to the latter maximum.

In summary, we have proposed and studied a mathematical model based on the assumption that matrix-binding of paracrine signals mediates early stages in *in vivo* vascular patterning. Simulation of the model suggests that the assumptions made are sufficient to generate vascular networks comparable to those observed during quail embryonic vascular development.

---

## DYNAMICS OF VEGF MATRIX-RETENTION IN VASCULAR NETWORK PATTERNING\*

---

### 9.1 INTRODUCTION

The assembly and patterning of new blood vessels plays a critical role in many processes, ranging from embryonic development to a variety of pathologies and successful tissue engineering<sup>256,314,315</sup>. Vascular endothelial growth factor (VEGF) is a major regulator of vascular morphogenesis that controls proliferation, survival, shape and migration of endothelial cells<sup>277,316–318</sup>. Yet, its activities are complex and context dependent and the role of VEGF in vascular network pattern formation is not fully understood.

The chemotactic response of endothelial cells to VEGF has been extensively studied. VEGF binding to cell surface tyrosine kinase VEGF receptor-2 (VEGFR2) activates, among others, multiple downstream pathways related to actin reorganisation, filopodia extension and invasive, protrusive behaviour<sup>316,317</sup>. In some cases, VEGF can induce cell migration to distant locations by acting as a long-range signal<sup>319–321</sup>. However, VEGF also functions as a short-range signal that controls branching and local network patterning<sup>281,292,322–324</sup>. Such a short-range signalling activity only takes place when VEGF splice variants are present that are able to bind to specific extracellular matrix components. Interestingly, matrix-bound VEGF has been shown to stimulate sustained signalling through VEGFR-2 with altered patterns of tyrosine activation and it seems to be a more effective initiator of migration in endothelial cells than soluble forms of VEGF<sup>227</sup>.

---

\* This chapter is based on the publication: Alvaro Köhn-Luque, Walter de Back, Yoshimi Yamaguchi, Kenji Yoshimura, Miguel A. Herrero and Takashi Miura, *Physical Biology*, 10:066007, 2013. Author contribution: Walter de Back conceived and designed the study together with Alvaro Köhn-Luque, Miguel A. Herrero and Takashi Miura. Walter de Back, Alvaro Köhn-Luque and Takashi Miura analyzed the experimental data. Model construction, simulation and data analysis performed by Walter de Back and Alvaro Köhn-Luque. Walter de Back and Alvaro Köhn-Luque wrote the paper.

It has been proposed that vascular network patterning is driven by the retention of VEGF by matrix molecules<sup>102,227,292,324</sup>. Moreover, it has been shown theoretically that binding of signalling molecules to extracellular matrix molecules can, in principle, generate fine-grained spatially restricted cues that may guide endothelial cell migration into network patterns<sup>102</sup>. Yet, surprisingly little is known about the spatial distribution of VEGF in the vascular microenvironment or the dynamics that establishes it. Moreover, it remains unclear how the VEGF distribution contributes to the formation of reticular vascular networks.

In this study, we shed light on these questions by performing a detailed investigation of VEGF dynamics in a simplified, controlled *in vitro* situation. We employed Human Umbilical Vascular Endothelial Cells (HUVECs) assay in Matrigel with exogenous administration of fluorescent-labelled VEGF<sub>165</sub>, an isoform with known affinity for matrix molecules. Spatial localization of VEGF and identification of VEGF retention molecules was performed using immunohistochemistry, confocal microscopy and RT-PCR. Biophysical characterisation of diffusion coefficients, binding/unbinding rates and decay kinetics was carried out by a combination of ELISA assays and Fluorescence Recovery After Photobleaching (FRAP) analysis.

We found that the initially homogeneously distributed VEGF accumulates in the extracellular space around cultured HUVECs within 60 minutes. FRAP analysis in the proximity of HUVECs shows that binding is a key mechanism of growth factor accumulation in those areas. Colocalization of VEGF with Heparan Sulphate Proteoglycans (HSPG) and fibronectin molecules suggests these matrix components as major candidates for VEGF retention. In line with the idea that cells play a role in VEGF retention, VEGF decay was found to decrease in a cell density-dependent manner.

Together, these findings provide support for the hypothesis that absorption of VEGF to the extracellular matrix occurs in a cell-dependent manner leading to the spatially restricted cues that guide endothelial cells into network patterns.

To investigate whether the observed VEGF retention dynamics can drive HUVECs into network patterns, a hybrid reaction-diffusion and cell-based mathematical model was constructed. Using measured concentrations and biophysical parameters, simulations results show the establishment of fine-grained distribution of bound VEGF and coalescence of cells into network patterns at realistic timescales, for different cell densities.

Taken together, our results suggest that cell-dependent retention of exogenous VEGF by matrix molecules around HUVECs can guide vascular assembly and patterning.

## 9.2 MATERIALS AND METHODS

### 9.2.1 Cells

HUVECs were purchased from Lonza Inc. and maintained in EGM-2 culture medium, also from Lonza Inc. Culture medium was changed every 2 days. Cells with less than 10 passages were used for experimentation.

### 9.2.2 Network formation assay

Matrigel solution (BD Biosciences) was prepared on ice. 50  $\mu$ l of Matrigel was spread in the center well of 35 mm glass-bottom dish (Matsunami Glass inc.). The dish was kept at room temperature for 10 minutes to allow the Matrigel to solidify. HUVECs in a different chamber was detached from dish using 0.1% Trypsin-EDTA solution (Nacalai Tesque inc). The cells were centrifuged, trypsin solution was removed and resuspended in 2 ml of EGM-2 culture medium, containing several growth factors supplements, including VEGF<sub>165</sub>. Cells were seeded on Matrigel-covered dish in various cell densities. The dish was cultivated 18 hours, and pattern formation was observed using inverted microscope (Nikon TMD or Nikon Eclipse). For VEGF-free condition, we prepared EGM-2 culture medium without VEGF supplement and use it for the experiment.

### 9.2.3 Preparation of fluorescent-labelled VEGF

Recombinant Human VEGF<sub>165</sub> protein was purchased from Peprotech Inc. 100  $\mu$ g of the protein were covalently conjugated with Alexa-488 fluorescent dye using microscale protein labeling kit according to the manufacturer's protocol (Molecular Probes Inc.). We use 4.7  $\mu$ l dye solution for labeling 100  $\mu$ g VEGF protein, and fluorescently-labelled protein was purified from unbound dye using gel filtration column according to the manufacturer's protocol. The labelled protein has a biological activity monitored by induction of dpERK in HUVECs (data not shown).

### 9.2.4 Absorption of fluorescent-labelled VEGF around HUVECs

HUVECs were seeded on thin layer of Matrigel on 35 mm glass-bottom dish as described above. The cells were cultivated for additional 10 hours at 37 °C in a humidified atmosphere. Then the dish was moved to stage-top incubator on Nikon C1 confocal microscope. 1  $\mu$ g/ml of fluorescent-labelled VEGF was prepared by mixing the EGM-2 culture medium with 1  $\mu$ g/ $\mu$ l stock solution. The working solution was prewarmed in the stage-top incubator. The culture medium in the culture dish was



completely removed, and fluorescent VEGF-containing culture medium was added to the culture dish. Then the absorption of fluorescent-labelled VEGF was observed at x60 magnification using Nikon C1 confocal microscope. Brightfield and fluorescence images were captured every minute.

#### 9.2.5 Fluorescence Recovery After Photobleaching (FRAP)

FRAP experiments in the Matrigel region were done using thin layer of Matrigel. First we mixed 1  $\mu\text{g}$  of fluorescent-labelled VEGF to 100  $\mu\text{l}$  of phenol-red free Matrigel (BD Biosciences) on ice. 3  $\mu\text{l}$  of labelled VEGF containing Matrigel solution was put on a slideglass, and covered by a coverglass of 18x18 mm size (Matsunami Glass Inc.). The Matrigel on the slide was allowed to solidify for 10 minutes at room temperature in a moisture chamber. Then the slide was put in the stagetop incubator, and the sample was photobleached at x20 lens. The bleach spot size was approximately 4000  $\mu\text{m}^2$ , and the recovery time is 60 frames with 10 second interval.

FRAP experiments in the proximity of cells were done using the HUVEC sample after absorption of fluorescent-labelled VEGF has reached equilibrium, which occurs in 60-90 minutes. The spot size was kept approximately 10  $\mu\text{m}^2$  and the recovery was observed for every 3 seconds for 200 frames.

#### 9.2.6 Immunohistochemistry

HUVECs were cultivated for 18 hours on a thin layer of Matrigel as described above. The cells were fixed in 100% Methanol for 10 minutes. The nonspecific binding was blocked using 1.5% NGS, and primary antibody against HSPG (Seikagaku Kogyo Inc. 10E4) or fibronectin (SC-69682 Santa Cruz biotechnology inc.) was added overnight at 4 °C. For HSPG, the culture dish was washed three times with PBS to remove unbound primary antibody. Then the biotin-conjugated secondary antibody was added to the dish, incubated for 30min at room temperature, and FITC-avidin was used to visualize the distribution. For fibronectin immunohistochemistry, we washed the dish, incubate in Alexa 488-labelled secondary antibody. In both cases, nuclei were stained with Hoechst 33342 to clarify the location of cells.

#### 9.2.7 RT-PCR

Total RNA of the HUVECs on the Matrigel was collected using Sepasol (Nacalai Tesque Inc.) The collected RNA was amplified by RT-PCR using RT-PCR High Plus (Toyobo Inc.) The primer pairs used were:

hFibronectin-F: AGAGAAGTGGTCCCTCGGCC

hFibronectin-R: TGGATTGAGCCCCGACCGT

hGlypican-F: TTCAGCCTGAGCGACGTGGT  
 hGlypican-R: TGCCCAGGCAGTCCAGGTAG  
 hPerlecan-F: CCTGGAGTGTGTCAGTGCCG  
 hPerlecan-R: CGAAGCGTGCTCTGGGACCG

#### 9.2.8 Sandwich ELISA

100  $\mu$ l of capture antibody were applied to each well of the 96-well plate and incubated overnight. Then each well were washed with wash buffer. Nonspecific binding was blocked by blocking buffer. After these preparations, dilutions of unknowns and standards were applied to the wells, and incubated for two hours at room temperature. Then the wells were washed with wash buffer. Biotynilated detection antibody was applied to the well and incubated for 2 hours at room temperature. After wash, 100  $\mu$ l of Streptavidin-HRP was added to each well and incubated for 20 minutes at room temperature. Then the substrate was washed away, 100  $\mu$ l of substrate solution (R&D systems) was added to each well and incubated for 20-30 minutes at room temperature to develop colour. Then 50  $\mu$ l of stop solution was added to each well. The optical density of each well was determined within 30 minutes with microplate reader (Thermo Fisher Multiskan FC).

#### 9.2.9 Mathematical model

A hybrid reaction-diffusion and cell-based model was constructed that accounts for key molecular and cellular mechanisms. The model structure, which is based on previous studies<sup>102,252</sup>, has been adapted to our particular *in vitro* experimental conditions. The model consists of two coupled modules. In the first one, HUVEC cells are represented as discrete and geometrically extended objects using a cellular Potts model (CPM)<sup>79</sup>. In the second one, growth factor and extracellular matrix molecules are modeled as continuous fields whose distribution is governed by partial differential equations (PDEs). Below, the main assumptions made for the formulation of the model are summarized:

1. HUVECs are seeded on top of a planar Matrigel surface and once they lie on it, their vertical displacement is small. Therefore, we reduce the problem to a 2D configuration.
2. Initially, the culture medium contains a homogeneous concentration of soluble, unbound VEGF. VEGF diffuses through the medium and the Matrigel and it can bind (and unbind) to various ECM molecules such as heparan sulphates

or fibronectin. For simplicity, we assume that such a reversible binding reaction follows the mass action law with second order kinetics.

3. The presence of HUVECs causes local accumulation of available VEGF binding sites of the surrounding extracellular matrix molecules such as fibronectin and heparan sulphates. This may be either through secretion of ECM molecules containing them, reallocation of existing ones, or opening of existing VEGF binding sites<sup>261,281,325–328</sup>. Accordingly, we assume accumulation of free VEGF binding sites to be cell-dependent.
4. The effective diffusivity of matrix molecules, with or without bound VEGF, are assumed to be very small compared to diffusivity of soluble VEGF, and are therefore neglected in the mathematical model.
5. Bound forms of VEGF provide stronger signalling cues for cell motility than freely diffusive forms<sup>227,281,292</sup>. Consequently, in the model motility is biased towards upward gradients of anchored growth factors, while chemotaxis towards freely diffusive forms is neglected.
6. Binding of VEGF to ECM molecules offers protection to growth factors against enzymatic degradation<sup>284,296</sup>. Therefore, we neglect degradation or removal of bound growth factors in the model.
7. No proliferation or cell death is taken into account during the modelled time window.

These assumptions lead to the following formulation of the PDE model. Let  $u$ ,  $s$  and  $b$  be the concentrations of unbound VEGF ( $[VEGF_u]$ ), specific VEGF binding sites of the ECM ( $[ECM_s]$ ) and bound VEGF ( $[VEGF_b]$ ) respectively. Bearing in mind assumptions (2), (3) and (5) above, we propose the following reaction-diffusion system:

$$\begin{aligned}
 \frac{\partial u}{\partial t} &= D_u \Delta u - k_{on} u s + k_{off} b - \tau u \\
 \frac{\partial s}{\partial t} &= \Gamma(\sigma_x) - k_{on} u s + k_{off} b \\
 \frac{\partial b}{\partial t} &= k_{on} u s - k_{off} b,
 \end{aligned} \tag{9.1}$$

where  $D_u$ ,  $k_{on}$ ,  $k_{off}$  and  $\tau$  represent the diffusion coefficient, binding, unbinding and decay rates of  $VEGF_u$ , respectively. The function  $\Gamma(\sigma_x)$  represents the rate of cell-dependent accumulation of free VEGF binding sites ( $\Gamma(\sigma_x) = 0$  in sites,  $x$ , that are not occupied by cells, while  $\Gamma(\sigma_x) = \gamma$  in sites,  $x$ , occupied by cells - see below).

Cells are represented on a discrete square lattice. The same index  $\sigma_x = 1, 2, \dots, N$  labels all the lattice sites,  $x$ , occupied by a particular cell, while the special index  $\sigma_x = 0$

labels all lattice sites,  $x$ , occupied by the extracellular medium. The interfaces between two different lattice sites,  $x$  and  $x'$  with different indexes  $\sigma_x$  and  $\sigma_{x'}$  represents the membrane boundaries between two cells or between a cell and the medium. Cell shape and position of cells are determined by volume and perimeter constraints plus a biased (chemotactic) movement towards bound VEGF. More precisely, we define the state of the system by means of the following Hamiltonian:

$$H = \sum_{\sigma > 0} (\lambda_a (a_\sigma - A)^2 + \lambda_p p_\sigma), \quad (9.2)$$

where the term  $(a_\sigma - A)$  represents the deviation of a cell labelled with the index  $\sigma$  from a target area  $A$  and  $p_\sigma$  is the total perimeter of cell  $\sigma$ . The parameters  $\lambda_a$  and  $\lambda_p$  represent the strength of the area and perimeter constraints respectively.

Dynamics in the discrete lattice are generated by a modified Metropolis algorithm that randomly chooses a lattice site,  $x$ , and computes what the difference in energy,  $\Delta H$ , will be if a randomly selected neighboring site,  $x'$ , would copy its state into this site. The probability of accepting the change,  $P$ , depends on the difference in the energy plus a chemotactic term  $\Delta H_c = \Delta H + \mu(b_x - b_{x'})$  measuring the difference in concentration  $b$  in the two lattice sites:

$$P(\Delta H_c) = \begin{cases} 1 & \Delta H_c \leq 0 \\ e^{-\Delta H_c} & \text{otherwise} \end{cases} \quad (9.3)$$

In this way, the shape and position of the cell are updated locally such that cell shape is biased towards a target area and minimum perimeter while their movement is biased up the gradients of  $VEGF_b$ . Note that cell adhesion has not been explicitly modelled.

The PDE and CPM modules are coupled by a feedback mechanism in which accumulation of free VEGF binding sites depends on the location of cells while cell shape and motility depends on local gradients of matrix-bound VEGF. While the PDE module has been parametrized with measurements of biophysical parameters and concentrations obtained in this study, the strength of the chemotactic response by CPM cells to VEGF gradients has been calibrated to reported chemotactic migration of HUVECs using published data.

### 9.2.10 Parameterization and calibration

The parametrization and calibration of the mathematical model described in section 9.2.9 are summarized in Table 9.1.

Parameters of VEGF dynamics in the PDE module are based on our experimental estimates. Specifically, the diffusion coefficient,  $D$ , binding and unbinding rates,  $k_{on}$  and  $k_{off}$  respectively, were estimated by means of FRAP as described in the main text.

Parameter	Estimate	Method
PDE		
D	$5.87 \cdot 10 \mu\text{m}^2\text{s}^{-1}$	FRAP
$k_{\text{on}}$	$8.57 \cdot 10^{-7} \text{ ng ml}^{-1}\text{s}^{-1}$	FRAP
$k_{\text{off}}$	$3.6 \cdot 10^{-3} \text{ s}^{-1}$	FRAP
$\epsilon$	$2.67 \cdot 10^{-6} \text{ s}^{-1}$	ELISA
$\gamma$	$0.5 - 2.0 \text{ ng ml}^{-1}\text{s}^{-1}$	Estimated in this study
CPM		
A	$300 \mu\text{m}^2$	Measured in this study
$\lambda_a$	$0.5 \text{ a.u. } \mu\text{m}^{-2}$	<sup>102,252</sup> rescaled
$\lambda_p$	$1.6 \text{ a.u. } \mu\text{m}^{-1}$	<sup>102,252</sup> rescaled
$\mu$	$40 - 140 \text{ a.u. ng}^{-1}\text{ml}$	Calibration using data from <sup>262</sup>
Initial condition		
$u_0$	$1.5 \text{ ng ml}^{-1}$	ELISA
$s_0$	$0.0 \text{ ng ml}^{-1}$	Assumed in this study
$b_0$	$0.0 \text{ ng ml}^{-1}$	Assumed in this study

Table 9.1.: Model parameters for equations 9.1, 9.2, and 9.3.

VEGF decay,  $\epsilon$  was measured using ELISA. As initial condition for VEGF,  $u_0$ , we use the actual concentration of VEGF present in the culture medium, while we assume there are initially no binding sites and no bound VEGF, i.e.,  $s_0 = b_0 = 0$ .

Cells are modeled phenomenologically as chemotactic objects with explicit shape using the cellular Potts model (CPM). In this formalism, cell shape as well as chemotactic motility arises from an abstract principle of energy minimization. The parameter values controlling cell shape elasticity,  $\lambda_a$  and  $\lambda_p$ , were adopted from previous studies on vascular patterning<sup>102,252</sup>, while the target area  $A$  is based on our own estimates on mean HUVEC cell size.

The key parameter coupling the CPM module to the PDE module is the chemotactic sensitivity  $\mu$ . This parameter relates the velocity of chemotactic migration of modeled cells to external gradients of VEGF. Fortunately, the chemotactic migration of HUVECs under influence of VEGF gradients has been studied in detail<sup>262</sup>. This allows us to calibrate the chemotactic response of our modeled cells to that observed in HUVECs under identical gradients of VEGF. Given choices of parameters for cell shape and area, the migration velocity is tuned to data on HUVEC migration by adjusting the chemotactic sensitivity  $\mu$  to obtain a chemotactic response that is similar to that of HUVECs.

Shamloo et al.<sup>262</sup> exposed a population of HUVECs seeded on fibronectin to a stable gradient of VEGF using a microfluidic device. They quantified the chemotactic sensitivity of HUVECs by measuring their migration under different gradients of VEGF. Migration was estimated by dividing the chamber into four equal zones of 250  $\mu\text{m}$  and measuring the net change in cell density before and after 6 hours of exposure to a gradient of VEGF. For a gradient of 14  $\text{ng ml}^{-1} \text{mm}^{-1}$  (with average absolute concentration 25  $\text{ng ml}^{-1}$ ), an average net change between 7% and 11% (+/- 3) in/out flux was reported for the two zones having the highest/lowest VEGF concentrations.

We replicated these experiments *in silico* by exposing the simulated cells to a static gradient of the chemoattractant of 14  $\text{ng ml}^{-1} \text{mm}^{-1}$  (replacing the PDE module) and measuring the change of cell density in four zones. By fitting the changes in cell densities in this model to the migration reported by Shamloo et al., we obtained an estimate for the chemotactic sensitivity parameter  $\mu \approx 3$ . Using this value for  $\mu$  in our model, cellular network can be obtained, but only by assuming a high production rate of VEGF binding sites (in the order of  $\approx 10^3 \text{ ng ml}^{-1} \text{s}^{-1}$ ). Under these conditions, the network structure, quantified by the number of lacunae, is established within 10 min, after which cells become prematurely immobilized due to the relatively weak chemotactic sensitivity (data not shown).

If, however, the chemotactic sensitivity is calibrated to a gradient of VEGF that takes matrix binding into account, different values for  $\mu$  are obtained. Although the gradient reported by Shamloo et al. accounts for the soluble VEGF in the microfluidic device, HUVECs migration is, we propose, controlled by the matrix-bound VEGF. Therefore, the gradient of matrix-bound VEGF is more suitable for the calibrating the chemotactic

sensitivity. Fortunately, the gradient of matrix-bound VEGF in the experiments by Shamloo et al. can be estimated. Given the quantified gradient of soluble VEGF and our measurements for binding/unbinding rates, the concentrations of bound VEGF can be calculated as the equilibrium of the following equations:

$$\frac{ds}{dt} = -k_{on}(14(x/L) + 18)s + k_{off}b \quad (9.4)$$

$$\frac{db}{dt} = k_{on}(14(x/L) + 18)s - k_{off}b, \quad (9.5)$$

where  $L$  is the length of the domain ( $L = 1000\mu\text{m}$ ). Here, the concentration of soluble VEGF,  $u$ , is assumed to be stable in the microfluidic device, and can therefore be replaced by the gradient.

Although we can assume the absence of bound VEGF before the onset of the VEGF gradient  $b(0) = 0$ , the initial condition is VEGF binding sites,  $s(0)$ , is not directly known. The authors used a concentration of fibronectin of  $10 \text{ mg ml}^{-1}$  in a hydrophobic PDMS situation under physiological pH (5%  $\text{CO}_2$  at  $37^\circ \text{C}$ ). Under these conditions, the number of binding sites per fibronectin molecule that VEGF can bind to, is reported to be approximately  $0.1 - 0.5$ <sup>327</sup>. This corresponds to a concentration of available VEGF binding sites of approx.  $75 - 375 \text{ ng ml}^{-1}$ . Using this range as initial conditions  $s(0)$ , simulations show the establishment of a gradient of  $0.25 - 1.25 \text{ ng ml}^{-1} \text{ s}^{-1}$ .

Using this gradient  $0.25 - 1.25 \text{ ng ml}^{-1} \text{ s}^{-1}$  of matrix-bound VEGF instead of the gradient of soluble VEGF to calibrate the chemotactic sensitivity of CPM cells, the sensitivity corresponding to the reported HUVEC migration rates becomes much higher,  $\mu = 40 - 140$ , see Figure 9.1 and the Supplementary Online Material. Using these values for chemotactic sensitivity parameter  $\mu$  in our model, a cellular network can be obtained under more reasonable values for production rate  $\gamma$ . Moreover, cells do not become prematurely immobilized.

### 9.2.11 Simulations

Numerical simulations were performed using our modelling software Morpheus. The modified Metropolis algorithm governing cell motility in the cellular Potts module (CPM) was specified with temperature  $T=1$  and a 8-pixel neighborhood. Random numbers were generated using Mersenne Twister algorithm (mt19937) available in C++ TR1 library extensions. Reactions in the PDE module were solved using the 4th order Runge-Kutta method with time step  $dt=1.0$ , i.e. once every Monte Carlo step. Space discretization was chosen equal for the CPM and PDE models, with each lattice sites corresponding to  $4 \mu\text{m}^2$ , and periodic boundary conditions were used. Diffusion was solved using Euler forward diffusion using time steps satisfying the Courant-Friedrichs-Lewy condition for the given length interval and diffusion coeffi-

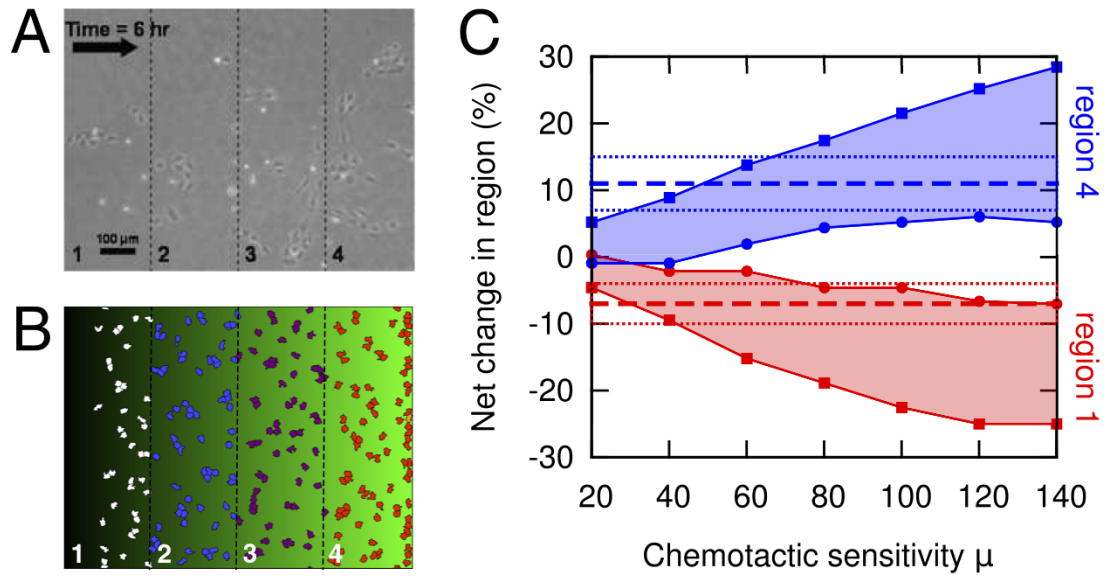


Figure 9.1.: **Calibration of chemotactic strength parameter.** (A) HUVECs migration in microfluidic device under stable VEGF gradient<sup>262</sup>. Migration was quantified by recording the change of occupancy in the indicated regions over 6 hours. (B) Computer simulation of chemotactic CPM cells under VEGF gradient, quantified in the same way. Cell colour indicates region, background colour shows gradient of bound VEGF. (C) Quantification of CPM cell migration under varying values for chemotactic sensitivity  $\mu$  by measured net change in regions 1 (red) and 4 (blue) (negligible in other regions). Filled areas show net change between the upper and lower bound for the estimated binding sites ( $375 \leq b \leq 75$ ). Dashed line and surrounding area shows average and standard deviation for the net change as reported in<sup>262</sup>. The model used for calibration is available in the Supplementary Online Material under <http://walter.deback.net/thesis>.



cient. The simulation model is available in the Supplementary Online Material under <http://walter.deback.net/thesis>.

### 9.3 RESULTS AND DISCUSSION

#### 9.3.1 Exogenous VEGF is required for network patterning

Network formation in our HUVEC cultures requires the administration of exogenous VEGF to the culture medium. We compared HUVEC cultures containing exogenous VEGF (control) with cultures with a VEGF-free medium. Quantification of number and total length of cell protrusions showed a marked decrease in the case of VEGF-free medium. Moreover, a dramatically reduced ability for network formation was observed in the VEGF-free case as compared with the network formed in the presence of VEGF. This indicates that other sources of the growth factor, such as the one present in Matrigel or that possibly produced by cells are not sufficient for network patterning.

#### 9.3.2 Exogenous VEGF is absorbed around cultivated HUVECs

In order to observe the spatio-temporal distribution of VEGF, rather than network pattern formation, we used an experimental setup with sufficient concentration of fluorescent-labelled VEGF<sub>165</sub>. More precisely, we cultivated HUVECs on a thin layer of Matrigel for 10 hours and applied 1 µg/ml solution of fluorescent-labelled VEGF on the culture dish. Surprisingly, accumulation of fluorescent-labelled VEGF around cells was observed within 5-10 minutes with fluorescence resulting in equilibrium after 60-90 minutes. Comparison of fluorescent and brightfield images shows that absorption is not only confined to the cell membrane, resulting from receptor binding, but also occurs at the extracellular region near the cells (figure 9.2).

#### 9.3.3 VEGF binding molecules colocalize with VEGF absorption areas

To investigate the biochemical reason behind the accumulation of fluorescent-labelled VEGF in the proximity of HUVECs, we labelled known VEGF binding molecules in our cultures by means of immunohistochemistry. We found that fibronectin and heparan sulfate proteoglycans (HSPG), two types of ECM molecules that strongly bind to VEGF, colocalize with VEGF absorption areas in the vicinity of HUVECs. It remains unclear to what extent the localization of ECM molecules is the result of secretion or redistribution of existing molecules by HUVECs. Although expression of both fibronectin and HSPG in HUVECs lying on a plastic dish was detected (data not shown), an active reorganisation of the existing VEGF binding molecules in Matrigel by cells cannot be excluded.

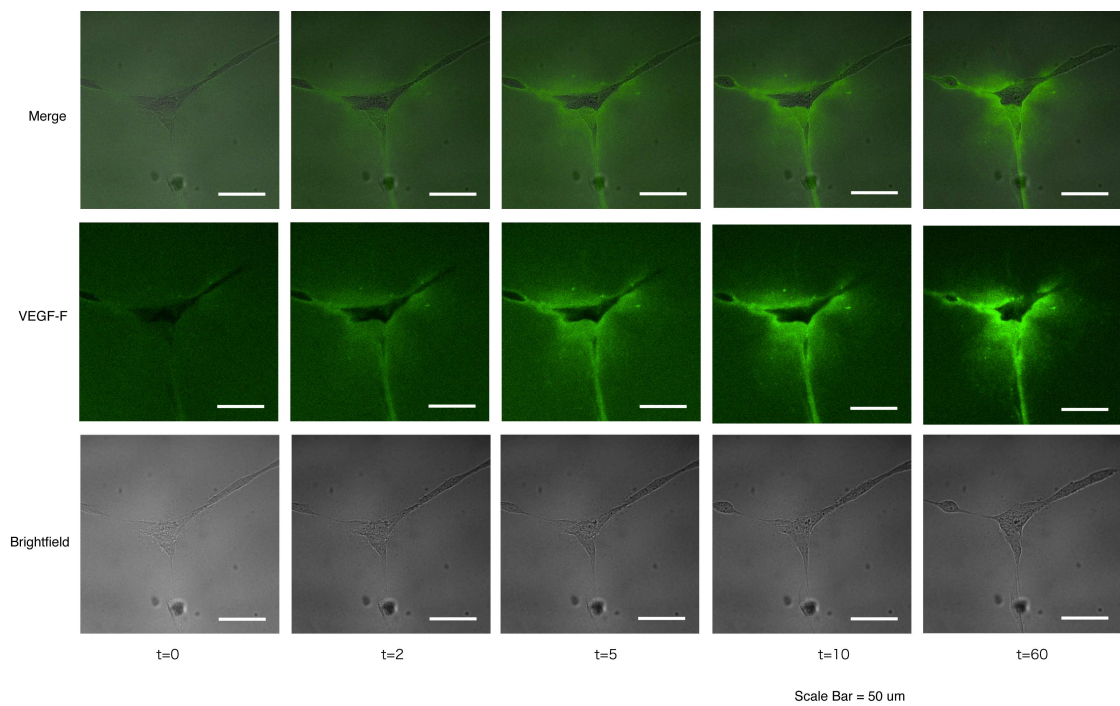


Figure 9.2.: **Absorption of VEGF around endothelial cells.** Time-lapse of fluorescent and brightfield images showing typical observed accumulation of fluorescently-labelled VEGF around HUVECs (Brightfield). A superposition of both image types is shown in the top line (Merge).

Nevertheless, the colocalization of the two ECM molecules, fibronectin and HSPG, and VEGF absorption areas together with their high affinity to VEGF render these molecules candidates for retention of VEGF in the proximity of HUVECs.

#### 9.3.4 Characterisation of VEGF<sub>165</sub> kinetics: FRAP analysis

To characterise the mechanisms of accumulation of VEGF<sub>165</sub> in pericellular regions, Fluorescence Recovery After Photobleaching (FRAP) analysis was performed. This allows us to distinguish between alternative mechanisms of growth factor accumulation and to estimate kinetic rates for transport and molecular interaction.

FRAP is a widely used technique to investigate transport and interactions of molecules<sup>329-332</sup>. In FRAP experiments, fluorescence tagged molecules are irreversibly photobleached in a small region of interest (ROI) by a high-powered focused laser beam. Subsequent displacement of surrounding non-bleached fluorescent molecules into the bleached area leads to a recovery of fluorescence, which is recorded at low laser power. By fitting experimental recovery curves to suitable theoretical models of VEGF dynamics, we can estimate the kinetic rates for the diffusion coefficient or binding/unbinding rates of fluorescent-labelled VEGF molecules present in Matrigel cultures.

First, we study VEGF diffusion in Matrigel in the absence of cells. A thin layer of Matrigel containing fluorescently-labelled VEGF is prepared. After the gel has solidified, a circular area of it is bleached and the recovery of the fluorescence in it is tracked until recovery saturates (figure 9.3A). If the increase in fluorescence is interpreted on the basis of a linear diffusion model, a formula for the recovery of the fluorescence can be derived<sup>333</sup>:

$$\text{frap}(t) = F_{\infty} \exp \left( -\frac{2}{1 + 8Dt/a^2} \right), \quad (9.6)$$

where  $F_{\infty}$  is the fluorescence after the recovery is complete,  $t$  is time and  $a$  is the radius of the bleached ROI. The unknown parameter  $D$ , the diffusion coefficient of VEGF, can be estimated by fitting eq. 9.6 to the observed recovery (figure 9.3, a4). Applying this procedure for three different recovery data, we obtained:

$$D \approx 5.87 \times 10^{-7} \text{ cm}^2/\text{s} \pm 2.1 \times 10^{-7}, \quad (9.7)$$

which is consistent with previous estimates<sup>278,334,335</sup>.

Next, we analysed the dynamics of exogenous VEGF in the areas around HUVECs where VEGF accumulation was observed. More precisely, we photobleached a small circular region around a cell (figure 9.3B), observed recovery of fluorescence and used models to fit it. In contrast to the cell free case, the linear diffusion model does not fit well to the recovery data obtained in the pericellular region. Further mathematical analysis shows that the observed diffusion of VEGF in the absence of cells (eq. 9.7)

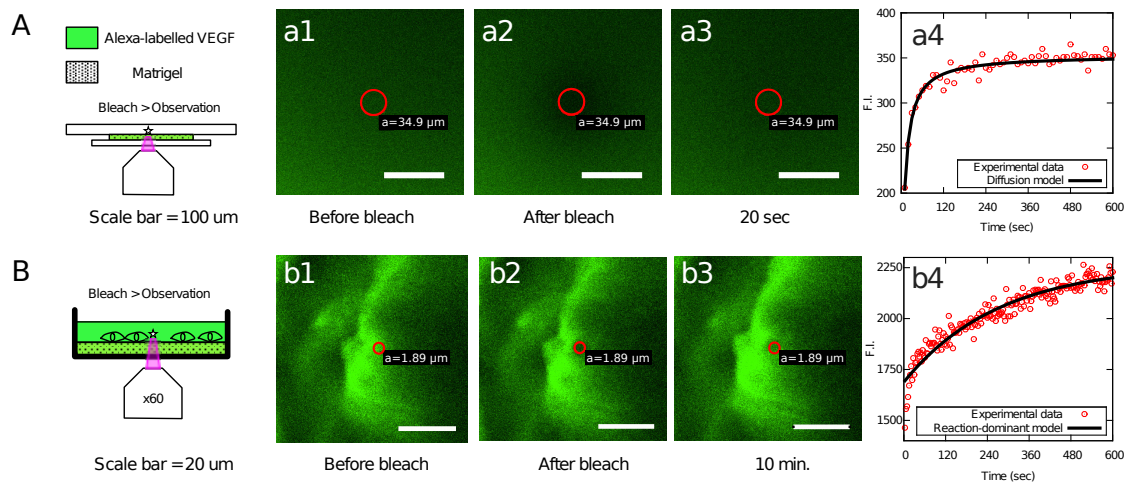
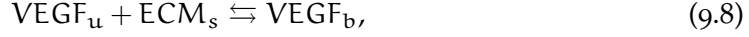


Figure 9.3.: **FRAP analysis.** Two different FRAP experiments are shown in top and bottom lines. (A) FRAP experiment in Matrigel in the absence of endothelial cells. Panels (a1), (a2) and (a3) show the observed VEGF fluorescence in three different time points, while (a4) shows a fitting of the linear diffusion model to the observed fluorescent intensity (F.I.), where the coefficient of determination corresponds to  $r^2 = 0.9328$ . (B) FRAP experiment in a culture of HUVECs in Matrigel. In this case the photobleached region is a small circumference in the proximity of an endothelial cell. Panels (b1), (b2), (b3) and (b4) are as in the previous case. The theoretical model used in this case to fit the obtained experimental recovery is the reaction-dominant model and the coefficient of determination is  $r^2 = 0.9999$ . Note that the spatial and temporal scales in both cases are different.

cannot be detected given the small size of the ROI and the time scale of the fluorescence measurements that are used in this experiment (see Appendix A.2 on page 161 for details).

Instead, based on the observation that the regions where VEGF is accumulated match the distribution of VEGF binding molecules of the ECM, we use a model that accounts for binding processes. A realistic case of binding that can be analysed by FRAP consist of a single binding interaction, described by the following chemical equation:



where  $\text{VEGF}_u$  and  $\text{ECM}_s$  denote unbound VEGF molecules and specific VEGF binding sites of the ECM, respectively.  $\text{VEGF}_b$  represents bound VEGF and  $k_{on}$ ,  $k_{off}$  are the binding and unbinding rates. In general, the equations describing the preceding binding reaction also incorporate diffusion (see Appendix A.3). However, we use here a particular case of the general model that adequately describes the recovery of the fluorescence in a tiny bleached area near a HUVEC cell. This particular case is commonly known in this context as reaction-dominant model and describes a scenario where VEGF diffusion is very fast compared both to binding and to the timescale of FRAP measurements<sup>331</sup>. In the reaction-dominant case, a formula for the fluorescence recovery can be obtained (see Appendix A.3 on page 165 for details):

$$\text{frap}(t) = u(t) + b(t) = 1 - \frac{k_{on}^*}{k_{on}^* + k_{off}} e^{-k_{off}t}, \quad (9.9)$$

where  $k_{on}^* = k_{on}S_{eq}$  and  $S_{eq}$  is an assumed equilibrium concentration of binding sites during the photobleaching recovery. Fitting eq. 9.9 to the FRAP recovery yields estimates for  $k_{on}^*$  and  $k_{off}$  (figure 9.3, b4). Applying this procedure for four independent recovery data sets, we obtained the following figures:

$$k_{on}^* \approx 1.5 \times 10^{-3} \text{s}^{-1} \pm 3.07 \times 10^{-4} \quad (9.10)$$

$$k_{off} \approx 3.6 \times 10^{-3} \text{s}^{-1} \pm 3.85 \times 10^{-4} \quad (9.11)$$

Obtaining  $k_{on}$  from  $k_{on}^*$  requires the equilibrium concentration of binding sites,  $S_{eq}$ , which is unknown. However, we can use estimated dissociation constants of VEGF to matrix molecules as fibronectin or HSPG, defined as  $K_d = k_{off}/k_{on}$ , that are available in the literature<sup>327</sup>. In that study,  $K_d$  has been shown to depend on pH and heparin concentration, but a suitable estimate for such parameter would be around  $100 \text{nM} = 10^{-7} \text{M}$ . Therefore,

$$k_{on} = \frac{k_{off}}{K_d} = 3.6 \times 10^4 \text{M}^{-1} \text{s}^{-1} \quad (9.12)$$

In summary, FRAP analysis shows that binding is a dominating mechanism behind the VEGF accumulation near HUVECs. Moreover, it provides quantitative estimates for kinetic parameters of  $\text{VEGF}_{165}$  diffusion as well as the binding and unbinding rates to ECM molecules.

### 9.3.5 VEGF is decreased in a cell density-dependent manner

ELISA assay was used to measure the evolution of VEGF concentration in culture medium alone (without Matrigel or cells) and in media extracted from cell cultures containing different cell densities. In the first case, VEGF decay is mainly due to molecular degradation. According to ELISA results, an estimate of VEGF half-life is above 72 hours. In the presence of cells, it was found that VEGF concentration in the culture medium is reduced in a cell density-dependent manner. Specifically, the VEGF concentration in the culture medium before and after 18 hours of HUVEC culture was measured. A significant reduction of VEGF was detected after culturing. The reduction was increased with higher cell density, indicating that HUVECs play an active role in the decrease of the concentration of soluble VEGF, possibly by uptake by cells or adsorption of VEGF to matrix components.

### 9.3.6 Model simulations mimic early stages of *in vitro* HUVEC cultures

To investigate the role of the observed VEGF retention dynamics in cell coalescence and network patterning, we performed computational simulations based on the measurements of kinetic rates and concentrations obtained in this study. Specifically, we modelled the early stages of the Matrigel cultures (approximately 0-2 hours), during which initially isolated cells coalesce and form network patterns.

Our hybrid mathematical model represents individual chemotactic cells as well as concentrations of VEGF and available VEGF binding sites. We assume the cell-dependent accumulation of VEGF binding sites and the binding of VEGF to those binding sites. In addition, we assume that bound forms of VEGF provide the spatial cues for HUVECs chemotactic migration. Model simulations were performed using our experimental measurements for kinetic parameters and concentrations of VEGF and the chemotactic response of simulated cells was calibrated to the chemotactic behaviour of HUVECs. See section 9.2.9 for a detailed description of model assumptions and formalization.

For these parameters, model simulations show the accumulation of VEGF in pericellular regions and the coalescence of initially isolated cells into reticular network patterns (figure 9.4), for a range of cell densities. Simulated networks are formed by cellular structures enclosing lacunae which occurs within two hours after administration of VEGF. During that period, the total number of lacunae rises and simulated cells increase their length up to three fold the initial one, correctly mimicking key features during early stages of our *in vitro* HUVEC cultures. Note that the coalescence and elongation of cells arises without explicitly modelling cellular adhesion and cell shape constraints. Rather, cells adjoin and elongate as a result of their chemotactic response to bound VEGF in the pericellular regions.

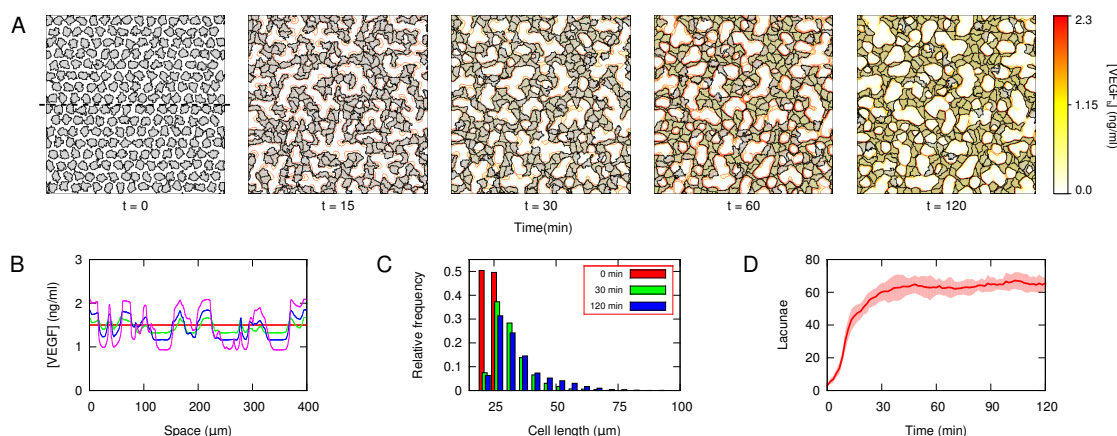


Figure 9.4.: **Network patterning in model simulations.** (A) Simulated cellular patterns (grey) and distribution of bound VEGF (colours) at  $t = 0, 15, 30, 60$ . (B) Redistribution of total VEGF concentration at the same time points, measured at the cross section indicated with dashed line in first panel A. (C) Distribution of cell length at  $t = 0$  (red),  $t = 30$  (green) and  $t = 120$  (blue) for simulation shown in A. (D) Evolution of number of lacunae, averaged over 10 simulations (average in red, standard deviation in light red region). The simulation model is available in the Supplementary Online Material under <http://walter.deback.net/thesis>.

These results show that vascular network formation is possible under the assumption of cell-dependent retention of exogenous VEGF, and occurs with the kinetics parameters obtained by FRAP and ELISA experiments.

## 9.4 CONCLUSIONS

Over the last decades, several hypotheses have been proposed to explain the formation of primitive vascular networks from isolated endothelial cells (see references<sup>241–244,336</sup> for recent reviews). Before the significance of VEGF in vascular patterning was recognised, theoretical work by Murray and collaborators showed that network patterns can arise solely due to matrix deformations caused by the traction that endothelial cells exert on the extracellular matrix<sup>248,337</sup>. This theory may explain why a number of cell types organise into network patterns when cultured in different extracellular matrices, including Matrigel<sup>338,339</sup>, but it does not elucidate the essential role of VEGF in vascular patterning. It might be that VEGF is merely required for cell survival and/or keeping endothelial cells in a patterning state. On the other hand, VEGF may be needed to spatially guide initial cell coalescence while mechanical forces become relevant only in a second stage, once sufficiently large cell densities are locally obtained<sup>311</sup>.

Related to this, alternative hypotheses for vascular patterning have focused in the role of VEGF as a chemoattractant for endothelial cells<sup>102,250,251,253,340</sup>. One such mecha-

nism holds that network patterning is based on the production of VEGF by endothelial cells themselves. Although such autocrine chemotaxis would typically lead to the formation of cellular clusters, several theoretical studies have shown that it can also lead to network patterning under certain conditions<sup>113,250,251,253,341</sup>. Yet, our result that the administration of exogenous VEGF is required for the formation of networks of HUVECs in Matrigel seems to be at odds with the autocrine mechanism. Indeed, studies have shown that during vasculogenesis in chick embryos, VEGF is expressed by immediately adjacent endodermal cells, whereas endothelial cells only express its receptor<sup>257</sup>. This strongly suggests a paracrine, rather than autocrine, activity of VEGF<sup>297</sup>.

Yet, the presence of a homogeneously distributed paracrine growth factor obscures, rather than explains, the formation of reticular network patterns. However, studies on mouse embryos that solely express a VEGF isoform without heparin-binding domains show an altered extracellular localization of VEGF and impaired vascular network formation<sup>292</sup>. This suggests that, under normal circumstances, VEGF binds to extracellular matrix components and thereby provides spatially restricted cues for endothelial cell migration<sup>292</sup>. Moreover, it was found that matrix-bound forms of VEGF elicit a distinct and prolonged signalling response in HUVECs in comparison to soluble VEGF<sup>227</sup>. Although these studies strongly suggest a role of matrix binding of paracrine VEGF, it has remained unclear how the spatial distribution of VEGF is established and how this affects the formation of cellular networks.

In the present study, we have shown that exogenous administrated VEGF quickly accumulates in pericellular regions, where we also found fibronectin and HSPG, two types of ECM molecules that bind strongly to VEGF. Moreover, a quantitative analysis by means of FRAP supports the idea that binding/unbinding is a dominating mechanism underlying VEGF pericellular accumulation. Although this process was recently hypothesized<sup>102</sup>, to the best of our knowledge it was not shown before. We have also provided estimates for binding/unbinding rates, diffusion coefficient and half-life of VEGF. Based on the previous knowledge, we have formulated a simple reaction-diffusion model for VEGF dynamics in the vascular microenvironment. In order to explore the role of bound VEGF in guiding endothelial cells independently of other factors or mechanisms, we couple the continuous model for VEGF dynamics with a cell-based model that accounts for chemotaxis towards bound VEGF of spatially extended simulated cells. The use of measured VEGF concentrations and experimentally derived parameters in the mathematical model enable us to interpret the simulation results in the correct spatial and temporal scales. In this way, we are able to conclude from our numerical simulations that the feedback between matrix retention of exogenous VEGF and chemotaxis alone can guide cellular coalescence and patterning in the early phases of vascular network formation. Our approach allows us to isolate the causal relations between matrix bound signals and network morphogenesis from a complex biological process influenced by many factors and difficult to manipulate.



Nevertheless, we do not underestimate additional factors others than VEGF such as, for instance, the juxtacrine signals that regulate multicellular sprouting<sup>263,269,302,342–344</sup>.

The combination of quantitative experiments and mathematical modelling employed in this work allows us to elucidate VEGF dynamics and its role in guiding vascular network formation. Apart from offering hypotheses to explain vasculogenesis in the embryo, we believe that this approach has promising applications in the design of engineered extracellular matrices for vascular development in tissue engineering and regenerative therapies<sup>345–348</sup>.



# 10

---

## CONCLUSIONS

---

Mathematical and computational methods are now available to study and analyze developmental systems in a formal, quantitative fashion. The interactions among genes, signals and cells can be represented in well-established modeling methods to study the regulation of cell fates, pattern formation and morphogenesis. The integration of these modeling methods into multiscale models facilitates the representation and analysis of regulatory feedbacks between these developmental mechanisms. These methods are important tools for the integration of quantitative biological information across temporal and spatial scales. It allows one to link the activity of regulatory molecules with the morphological development of organisms and enables the exploration and identification of self-organizing principles underlying a broad range of developmental phenomena.

In this thesis, I have described a modeling framework that adopts the cell as a basic level of abstraction to represent that the behavioral and mechanical interactions between cells within embryonic tissues. Taking a middle-out approach, the associated cell-based models can be readily integrated with models of intracellular regulatory dynamics and extracellular reaction-diffusion dynamics. The computational methods required for multicellular and multiscale modeling and simulation are available in number of recently released dedicated software platforms, including our modeling environment Morpheus. Morpheus is specifically designed for multicellular systems biology and is unique in its user-friendly design that allows researchers, even without computational expertise, to develop multiscale simulation models to test specific hypotheses about multicellular and developmental systems.

Two case studies have been presented that illustrate the application of this approach to various developmental systems. In the first, the regulation of cell fates and the emergence of fine-grained spatial patterns in the pancreas was investigated. A model was constructed that combines reported interactions between transcription factors with two contact-mediated mechanisms of cell-cell communication: lateral inhibition and lateral stabilization. Using bifurcation analysis, bistability of two spatial patterning

solutions is shown. Stochastic lattice simulations were used to demonstrate that this bistability can result in the scattered distributions of cell with the primary cell fate and that this allows for the modulation of cell type ratios. The results from chapter 5 shows that tissue-level phenomena such as the propagation of inductive waves provide a level of control over cell fates, spatial patterns and cell type ratios that cannot be understood from a single-cell perspective. These mechanisms can be exploited for cellular reprogramming, as shown in chapter 6. In cases where multistability depends on contact-mediated signals, selective perturbation of intercellular communication can be used to induce lineage conversion, without the need for genetic overexpression.

This case study demonstrates that the coupling between dynamic models of gene regulatory networks with cell-based models of cell-cell interactions provides useful tools to understand cell fate decisions as well as their spatial distribution. Moreover, it shows that spatial coupling between cells is an important determinant of cell fates and can be used to direct cell fates.

These results have already been used in a review on the role of models to understand the molecular regulation of pancreatic  $\beta$ -cell mass<sup>350</sup> and within a wider theoretical analysis of cell-cell communication through Notch-Delta-Jagged signaling<sup>159,160</sup>. In addition, these results can have various practical applications for stem cell biology and cellular reprogramming and recent reports in stem cell biology show exciting new results in this direction. For instance, it was shown that the initial seeding density influences the cell fate specification of human embryonic stem (ES) cells<sup>351</sup>; that controlling spatial structure of mouse ES cell colonies by micropatterning suffices to trigger self-organized patterning into three germ lines<sup>352</sup>; and that mouse ES cells colonies, depending on the correct size of the initial aggregate, can undergo spontaneous symmetry breaking and self-organize an axial organization and germ layers<sup>16</sup>. These recent observations are still poorly understood and more experimental and theoretical work is required. However, their explanation will most likely involve a combination of genetic regulation and cell-cell communication. Therefore, the methods described in this thesis provide a suitable theoretical framework to understand these phenomena. In a recent proof-of-principle study, we have already begun to apply these tools to investigate spatial expression patterns in mouse ES cells<sup>119</sup>.

The second case study focused on the morphogenesis of isolated endothelial cells into connected vascular capillary networks. This was modeled by combining a cell-based model of chemotactic cell migration with a reaction-diffusion system representing the diffusion and binding of a chemoattractant, VEGF, to the extracellular matrix. Even though the chemoattractant was assumed to be produced homogeneously in nearby tissues, computational simulation in chapter 8 showed that fine-grained network patterns arise for a wide range of parameters, when binding to cell-modified extracellular matrix molecules was assumed. The simulated networks show a striking similarity to *in vivo* images of extraembryonic vascular plexus in quail embryos, as

shows by quantitative morphometric analysis. Additionally, analysis of the dynamics of network formation in our model showed that after initial coalescence, substantial network remodeling occurs and that cell elongation is an emergent effect rather than a driving force of network formation, both contrasting results in previous models. In chapter 9, the paracrine model of vascular network formation was validated using a quantitative analysis of the dynamics of VEGF binding in an *in vitro* assay of HUVEC in Matrigel. As predicted, accumulation of fluorescent VEGF was observed in pericellular regions. Moreover, FRAP analysis showed that VEGF dynamics in those areas cannot be explained by diffusion, but fit well with a reaction-dominant that was derived from the model. Quantitative measurements on VEGF diffusivity, decay, matrix binding and unbinding were obtained experimentally and used for computational modeling. Simulation of the quantitative model confirmed the ability of network formation in the paracrine model, under biophysically realistic parameters.

In this study, the model demonstrated that a spatially homogeneous signal can result in the establishment of complex tissue architectures. It highlights the role of the extracellular matrix in this process as a storage for biochemical signaling factors apart from its function to generate and direct biomechanical forces<sup>353</sup>, for which there is more attention recently. The results of this case study have been recognized in a number of subsequent computational studies and reviews on blood vessel development<sup>242,244,249,264,336,340,354–358</sup>. Moreover, the result that matrix-retention of paracrine chemotactic signals is a key regulator of vascular network formation can be relevant for tissue engineering. The design of functional tissues and organs *in vitro* has so far been limited to thin tissues, because vascularization remains the main obstacle to obtain more complex functional tissues<sup>315,359,360</sup>. One of the key approaches to promote vascularization of engineered tissue is through the design of functionalized biomaterials that mimic biochemical properties of the extracellular matrix<sup>359,361</sup>. Instead of functionalization with VEGF directly, our modeling work rather suggests enrichment of VEGF binding sites, together with means for cells to induce VEGF signaling from the matrix-stored VEGF such as addition of RGD peptides that are known to stimulate VEGF signaling through interaction with integrins<sup>266,362</sup>. More specific suggestions, however, would require detailed data-driven modeling of the molecular interactions at the membrane-matrix interface, which is still outstanding.

The models used in these case studies are deliberately kept simple. In both studies, the system was reduced to two spatial dimensions, involved only two levels of organization, ignored important aspects such as tissue growth and proliferation and conveniently neglected possible interactions between induction and morphogenesis. Nevertheless, these modeling studies embody the key modeling methodologies of multicellular systems biology of development: (1) Integration of processes and modeling formalisms at multiple levels of biological organization, using the cell as the fundamental unit of abstraction. (2) Establishment of dynamic multiscale models from

piecemeal and descriptive experimental data. (3) Progression from theoretical first-principle studies towards quantitative modeling of developmental systems. (4) Iteration of computational modeling and experimental validation to establish quantitative models. Moreover, they have demonstrated novel uses of several key techniques in addition to multiscale model simulation: (1) Bifurcation analysis to examine the stability of tissue-level patterns and thereby to investigate how cell fate stability depends on cell-cell communication. (2) *In silico* experimentation to explore systems behavior and generate predictions. (3) Quantification of computational models by *in vitro* measurements and calibration of unknown model parameters by replicating published quantitative experimental studies. (4) Image-based morphometric analysis of network morphologies to assess the similarity of simulation results and *in vivo* observations.

Importantly, model simplicity is not a weakness, but an asset. Although the new methodological and technological tools in multicellular systems biology enable the construction of large-scale integrative quantitative models, whether such inclusive models generate new biological insights is doubtful. The power of modeling lies in exclusion and abstraction rather than inclusion and realism. The aim of modeling is to find the minimal set of mechanisms at the relevant levels of organization to explain a particular phenomenon. Thus, paradoxically, reduction is of key importance in integrative modeling.

The middle-out modeling strategy, combined with the cell-centered perspective, provide the guidelines for this reduction. Even at the cellular level, aspects such as cell shape, motility and intercellular communication are only explicitly modeled when required for the mechanism under investigation, or to explore their possible effects on the biological phenomenon. Models of processes at lower and higher levels, e.g. intracellular regulatory dynamics or extracellular morphogens, are coupled to the cellular level insofar as they are hypothesized to affect the behavioral or mechanical properties of the cell. Whether or not to include an additional process or submodel should not be motivated by presumed improvements in realism. Instead, the criterion is whether the extended model increases its ability to be used to explore the consequences of new interactions. Thus, the quality of a model is not related to its realism, but to its usefulness, as aptly phrased by George Box above<sup>349</sup>.

#### FUTURE PERSPECTIVES

Multicellular systems biology of development is an emerging field and its methodologies and techniques, as presented in this thesis, offer ample room for consolidation and improvement. Here, I highlight three directions and challenges for future research that I deem important to foster the maturation of this field of study.

## Open standards in multicellular systems biology

With the increasing complexity of computational models, the transparency, reproducibility and exchange of simulation models is a growing concern. In molecular systems biology, the definition of free and open standard exchange formats to represent biochemical networks such as SBML (Systems Biology Markup Language<sup>93</sup>) have boosted research over the last decade. It has facilitated model exchange between software as well as users, and the establishment of repositories of such standardized models (i.e. EBI Biomodels database) have provided a platform for knowledge transfer and promoted incremental progress in modeling biochemical networks.

Unfortunately, no standard exchange format exists for multicellular systems biology. Therefore, current models of multicellular systems are limited to the specific software platform in which they are developed. This obstructs reproducibility of published results, comparison of results obtained in different formalisms and software, and the free exchange of models and knowledge between researchers. The recent publication of a Cell Behavior Ontology<sup>363</sup> takes a first step by defining a structured language to describe cellular behaviors. However, it does not provide a way to translate these behaviors into computational models, let alone a conversion between different computational models and software platforms. And so far, none of the software platforms discussed in chapter 2 supports this ontology.

In a recent paper<sup>94</sup>, we identified that one of the key challenges in the standardization of multicellular models is the fact that, in most software platforms, computational models are implemented using one of several procedural programming languages (e.g. C++, Java, Fortran, Python) rather than described in a declarative markup language (e.g. XML). Yet, such a separation of a model from its implementation, i.e. abstraction of *what* a model does from *how* it is simulated, is a necessary condition to define exchangeable and platform-independent model descriptions<sup>118</sup>.

With the formulation of MorpheusML, the model description language used in Morpheus (see section 1), we have shown that it is possible to define a declarative domain-specific language specifically designed to flexibly describe multiscale multicellular models. Within Morpheus, this declarative language provides the additional benefits of enabling model editing in a graphical user interface without programming, facilitating model sharing between users, allowing automatic validation of model consistency and enabling semantic model integration. Structurally, MorpheusML is similar to SBML, as indicated by the fact that SBML models can be converted automatically. Therefore, MorpheusML can provide a prime example for the future development of a standard exchange format for multicellular systems biology.

## Image-based modeling

Microscopy images are the predominant source of data for multicellular systems biology and provide both the needs and the means for computational modeling. Recent advances in multichannel, 3D time-lapse imaging and automated, quantitative image analysis are rapidly delivering high-quality quantitative data on multiple levels of organization. There are a number of key challenges for the effectively use of image data to provide new insight into developmental processes.

Multicellular modeling increasingly depends on quantitative imaging in several ways. Quantitative image analysis is used to constrain model parameters to realistic values, i.e. cell division rates can be computed from BrdU labeling and, as in chapter 9, reaction-diffusion parameters may be obtained from FRAP analysis. Segmented image data can also be used directly to define initial or boundary conditions of objects in a computational models. And, importantly, validation of computation models often relies on quantitative image analysis, such as the direct morphometric comparison of simulation results with experimental observations, as in chapter 8.

Despite the increasing reliance of multicellular models on images and image-based quantitative data, there is a lack of methods and software tools that facilitate this pipeline. Therefore, establishing new tools for image-based modeling is a key topic for future development. As an initial step towards the integration with image analysis software, we have enabled the import of segmented TIFF images in Morpheus to define boundary conditions or initial cell configurations. Additionally, the export of simulation results in the form of TIFF image stacks for *post hoc* analysis enables the direct comparison of images obtain *in vitro* and *in silico* using the same pipeline for quantitative image analysis, as we have recently demonstrated in a study on spatial heterogeneity of embryonic stem cells<sup>119</sup>.

Although common data formats are important, a tighter integration scheme in which image analysis can be conducted during model simulation will open up various new opportunities such as efficient optimization of model parameters for 'fitting' model behavior to still or time-lapse images.

A less conventional but interesting combination of image analysis and modeling lies in the *in silico* generation of ground truth data sets to assess the quality of image analysis methods. Given the growing complexity of image processing algorithms, the need to measure and compare their performance on complex image data become increasingly dominant. Although a number of benchmark problems have been recently established<sup>364</sup>, these are mostly 2D problems and largely based on manual segmentation and tracking. This approach is insufficient to measure the performance of algorithms of 3D imaging of large specimen<sup>21,365,366</sup>, for which researchers are now resorting to computer-generated volume renderings to act as ground truth data sets<sup>366</sup>. While this works well for static 3D images, it cannot be used to generate artificial time-lapse



data sets that require a model for cellular behavior. Computational modeling of multicellular systems can provide helpful tools to generate artificial image data of tissue development through time for which the underlying mechanisms and parameters are completely known. These can then be used as ground truth data to assess the quality, in terms of sensitivity and accuracy, of the image analysis algorithms and statistical analysis.

## Morphodynamics

As outlined in the introduction (chapter 1), one of the major current challenges in developmental biology is understanding how cell fate specification, patterning and morphogenesis interact to give rise to the self-organization of tissues. As this involves simultaneously taking into account the dynamic interactions between genetic, chemical and mechanical signals, multiscale computational modeling is an invaluable method to formulate hypotheses and provide testable predictions on such interactions.

In the case studies presented in this thesis, these interactions have been investigated only in part. For reasons of simplicity and tractability, cell motility was largely ignored in the study on pancreatic cell fate decisions (with the exception of section 5.3.5), and the effects of cell fate differentiation and maturation were neglected in the study on vascular morphogenesis. However, the modeling framework as presented in chapter 2 as well as its implementation in the modeling environment *Morpheus* (chapter 3) readily support such computational studies in so-called morphodynamics<sup>12,27,28</sup>.

Indeed, there is a growing number of studies that use computational simulation to investigate complex developmental processes by explicitly accounting for dynamical interactions between genetic, chemical and mechanical processes at the molecular, cellular and tissue level. These works already cover a wide range of problems in developmental biology, such as the effects of cell packing and mechanical deformations on planar cell polarity<sup>104,367</sup>; the effect of cell motility on phase synchronization during segmentation<sup>14,368</sup>; the interactions between lateral inhibition, apoptosis and adhesion during mosaic patterning<sup>369</sup>; the role of cell rearrangements on sprouting angiogenesis<sup>370</sup>; the interaction between gene regulation and tissue mechanics in tooth development<sup>371</sup>; the feedback between hormone transport and tissue growth during phyllotactic patterning in plants<sup>372</sup>; and the interplay between mechanics, directed cell migration, cell lineage decisions and clonal competition in the formation of intestinal crypt organoids<sup>373,374</sup>.

Undoubtedly, the coming years will show a fast increase in the use of this type of multiscale computational modeling now that enabling technologies such as dynamic imaging, quantitative image analysis and multiscale simulation are becoming widely available. Regardless whether it will proceed under the name of quantitative developmental biology<sup>26</sup>, computational morphodynamics<sup>27,28</sup>, cytosystems dynamics<sup>11</sup> or multicellular systems biology of development, these studies will help to elucidate how

the network of interactions between genes, signals and cells gives rise to the self-organizing regulation of tissue development. It is my hope that the studies, methods and software presented in this thesis will contribute to this endeavor.

---

## ACKNOWLEDGEMENTS

---

I would like to express my gratitude to Andreas Deutsch and Lutz Brusch for their support, guidance and encouragement. I thank the members of the IMC group for the many constructive comments and for creating a stimulating academic environment.

In particular, I would like to thank Jörn Starruß for the intense and fruitful collaboration in the development of Morpheus. I also wish to thank my office mate Fabian Rost for the positive working atmosphere and the many useful advices. A special thanks is extended to Alvaro Köhn-Luque for our successful cooperation and all the enjoyable interactions, inside and outside of the office.

Finally, I thank my wife Christiane for her love and encouragement and, last but not least, I thank my children Bojan, Rosalie and Meta for offering many welcome distractions.



---

## APPENDICES

---

### A.1 COMPUTATIONAL PERFORMANCE AND SCALABILITY

The performance and scalability of simulations in Morpheus heavily depend on the type of (multi-scale) model that is being simulated. It is therefore difficult to make general statements on the computational efficiency. However, we can test the performance on a set of “benchmark” models that form the modules from which more complex model can be constructed.

We have tested the performance of ODE lattices, reaction-diffusion (PDE) models, cellular Potts models (CPM) and a multiscale model (CPM + PDE), using the available Example models. The results show the execution time and memory consumption for these models as well as their scalability in terms of problem size and scalability in terms of efficiency of multi-threading.

#### Methods


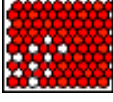
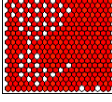



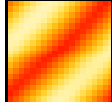
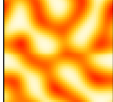







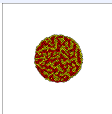
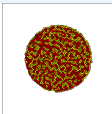
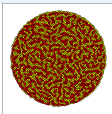
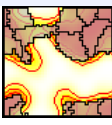
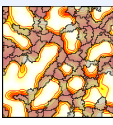
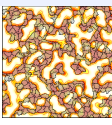
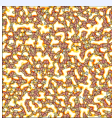
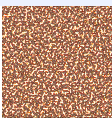
##### *Measurements*

To quantify performance, we measured the following aspects for each simulation. Execution time in terms of the wall time, using the C++ function `gettimeofday()` available in `<sys/time.h>`. The execution time does not include the time needed for initialization, analysis and visualization. Memory usage in terms of the physical memory (RAM) used by the simulation, using the resident set size (RSS) from the `/proc/self/stat` pseudo-file.

##### *Simulations*

The models and the variation of problem sizes are shown in table A.1. For ODE models, not the number of variables, but the domain size (=number of cells) is varied. For PDE models, the domain sizes is increased. For CPM simulations, the number of cells is varied, keeping a constant domain size. Note that edgelist tracking is used, such that the number of updates scales with total size of cell boundary instead of lattice size. For

Table A.1.: Simulations used for performance and scalability tests.

<b>ODE</b>						
Cells	25	100	400	2500	10000	40000
<b>PDE</b>						
Lattice	$20^2$	$50^2$	$100^2$	$200^2$	$500^2$	$1000^2$
<b>CPM</b>						
Cells	100	200	400	1000	2000	4000
<b>CPM+PDE</b>						
Lattice	$40^2$	$100^2$	$200^2$	$400^2$	$1000^2$	
Cells	8	50	200	800	5000	

CPM+PDE models, both domain size and cell number are increased proportionally, such that the cell density remains constant.

All models are simulated without analysis and visualization tools and execution time is measured from `StartTime` to `StopTime`. The time for initialization is excluded because this vanishes for large jobs. All simulations were performed on a Intel Core i7-860 vPro with 4 cores, 8 threads with (hyperthreading), 2.8 GHz clock speed, 8 MB cache and 20 GB RAM memory.

## Results

### *Scalability with problem size*

We investigated the scalability with respect to problem size to see how performance in terms of the execution time and memory usage (RAM) scales with increasing population size or lattice size. We calculate both the execution time and memory usage in absolute and relative terms. In absolute terms, the time is recorded in seconds (sec) and memory in megabytes (MB). In relative terms, the time and memory is divided by the number of cells or lattice site in millisecond (msec) / kilobyte (kB). Performance in absolute sense gives a sense of the problems sizes that are practically manageable

within certain time and memory constraints. Performance in relative sense shows the scalability of the simulation for problem sizes. Ideally, the performance per cell or lattice site stays constant or decreases with increasing problem sizes.

The results are shown in table A.2. The left column shows that absolute execution time scales approximately linearly with problem size, as expected. Moreover, the absolute memory usage is not a limiting factor in any of the tested simulations. Even the maximum total memory usage, found for the largest CPM+PDE simulation (5000 cells,  $1 \cdot 10^6$  lattice sizes), is well below the 200 Mb. The right column with results, relative to the domain size or cell number, shows that for the ODE and PDE simulations, the execution time per cell or lattice site does not increase with problem size, thus demonstrating perfect scalability. The CPM and CPM+PDE simulations, however, show a small increase in execution time per cell or lattice size for large simulations, suggesting a poorer scalability for large simulations. This can be explained by a more than proportional increase in cell-cell interface length in larger simulations, requiring more computations.

#### *Efficiency of parallel processing*

We have also measured the scalability with respect to the number of openMP threads to see how the performance scale with the number of concurrent threads. We measured the execution time for each of the simulation run on in 1, 2, 4, 6 threads. Comparison of these execution times shows the speed-ups that can be achieved by adding concurrent threads.

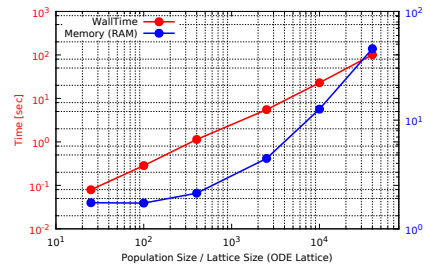
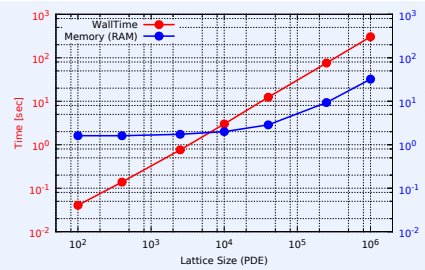
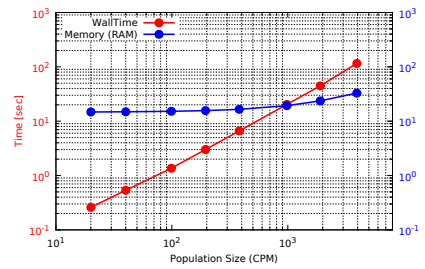
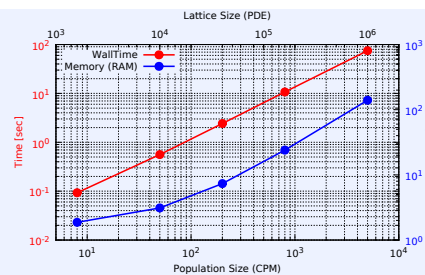
The results, in table A.3, show that PDE simulation are the most efficiently parallelized. For 2 threads, simulations are almost twice as fast. However, using 6 threads, simulations are only 3.5 times faster. This may be related to the fact that only the reaction step is parallelized, whereas the diffusion step is not. ODE simulations have only moderate benefit from multithreading. This is related to the fact that exchanging values from/to neighboring cells, rather than the computation of the intracellular ODEs, is a bottleneck in these simulations. The CPM simulations show no speed-up at all, because CPM simulations are not parallelized. In fact, a small slow-down is shown, due to the overhead of multithreading. Finally, the CPM+PDE simulations show only moderate speed-up, which combines the presence resp. absence of benefits of parallel computing in PDE and CPM simulations.

## A.2 OBSERVATION LIMIT FOR A FRAP EXPERIMENT

An initial homogenous distribution of exogenous VEGF was observed to accumulate around HUVECs. To investigate the dynamics of exogenous VEGF in those areas close to the cell membrane, a small circular ROI (approximately  $10 \mu\text{m}^2$ ) was photobleached and the recovery of the fluorescence was observed every 3 seconds. This Appendix

**Problem size - absolute**

Total execution time (red) and memory usage (blue) of simulation, excl. initialization and visualization

**ODE****PDE****CPM****CPM + PDE****Problem size - relative**

Execution time (red) and memory usage (blue), relative to number of cells and/or lattice sites

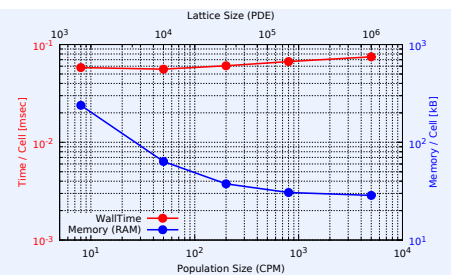
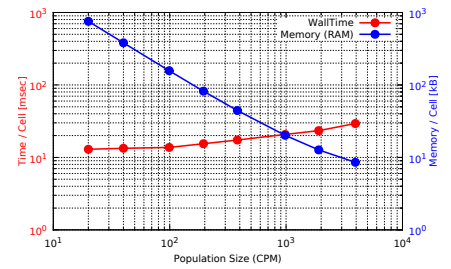
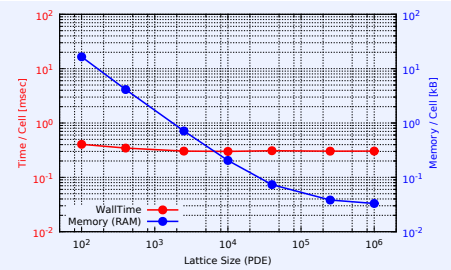
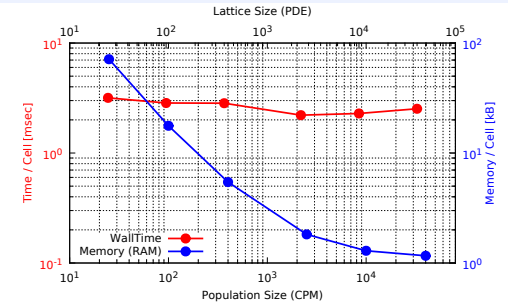


Table A.2.: Performance results for varying problem sizes.



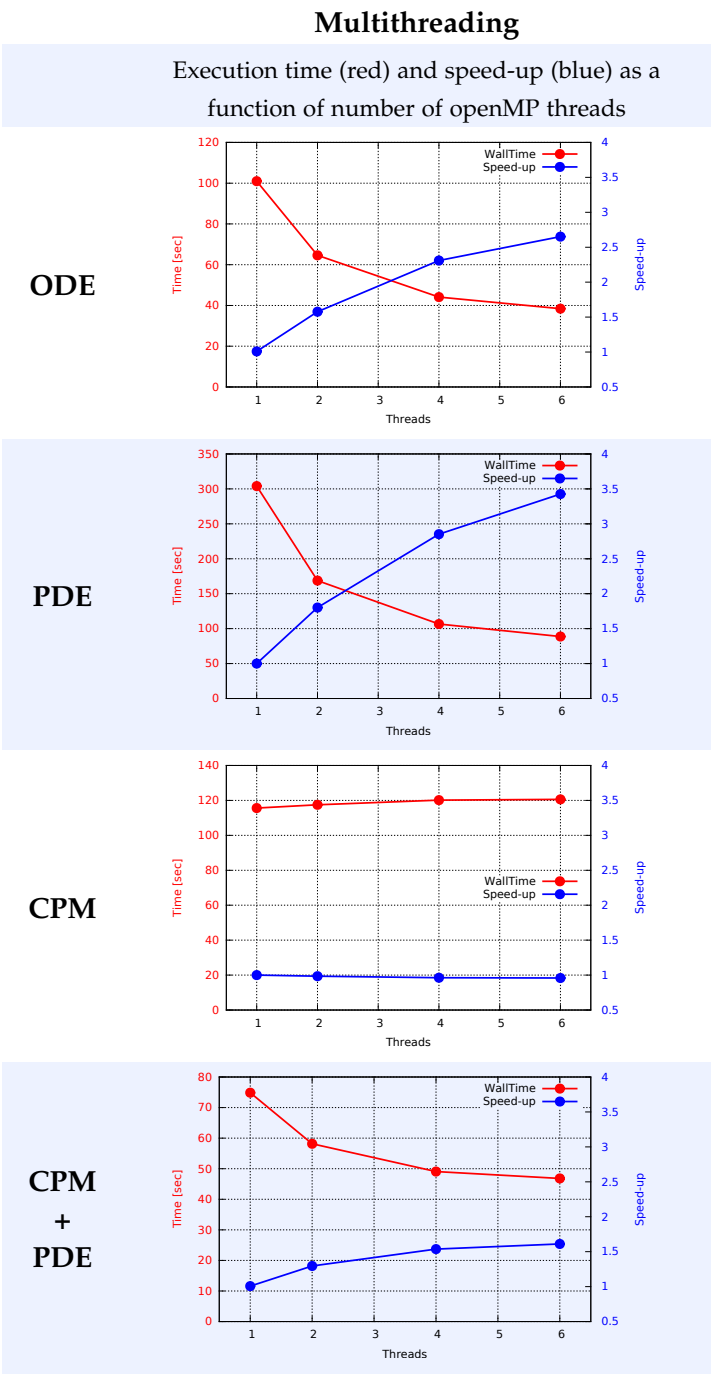


Table A.3.: Scalability with varying number of threads.

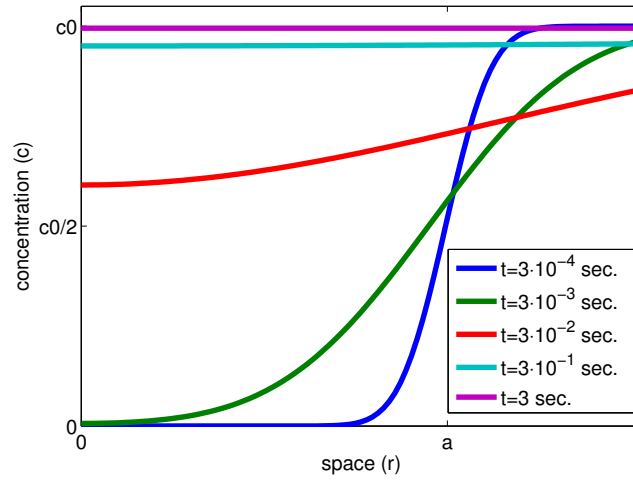


Figure A.1.: **Diffusion in a circular region after photobleaching.** The evolution in time of the concentration of fluorescent VEGF is shown with respect to the radial distance from the optical axis. After photobleaching, the concentration profile inside the circular region ( $r < a$ ) increases from zero to  $c_0$  in 3 seconds.

shows that under these experimental conditions, the estimated diffusion of VEGF in Matrigel, (eq. 9.7), is too fast to be detected. To do that, one can use an analytical formula for the concentration of the fluorescent molecule after photobleaching. Then, one can substitute therein the particular size and the estimated diffusion coefficient of VEGF and look at the time evolution of the concentration in the time scale of the experiment. More precisely, the concentration,  $u$ , of the fluorescent molecule is given by Jacobson's formula<sup>333</sup>:

$$u(r, t) = u_0 \left( 1 - \frac{\exp(-r^2/4Dt)}{2Dt} \int_0^a \exp(-r'^2/4Dt) I_0 \left( \frac{rr'}{2Dt} \right) r' dr' \right), \quad (\text{A.1})$$

where  $D$  is the diffusion of the molecule,  $r$  is the distance from the optical axis,  $t$  is time,  $a$  is the radius of the bleached ROI,  $u_0$  is the initial concentration for all  $r > a$ , and  $I_0$  is the modified Bessel function of the first kind of order zero. Figure A.1 shows the time evolution of the concentration  $u$  in eq. A.1 using the estimated diffusion coefficient shown in eq. 9.7, and the size of the ROI,  $a$ , used in figure 9.3B. Note that already after 3 seconds, the time interval between measurements in this FRAP experiment, the concentration is almost uniform and very close to the concentration of the fluorescence outside the ROI,  $u_0$ . Therefore, the recovery due to diffusion of free VEGF is not captured in this experiment.

## A.3 DERIVATION OF THE FRAP RECOVERY CURVE

In general, the equations describing the binding reaction (eq. 9.1) also incorporate diffusion:

$$\frac{\partial u}{\partial t} = D_u \Delta u - k_{on} u s + k_{off} b \quad (A.2)$$

$$\frac{\partial s}{\partial t} = D_s \Delta s - k_{on} u s + k_{off} b \quad (A.3)$$

$$\frac{\partial b}{\partial t} = D_b \Delta b + k_{on} u s - k_{off} b \quad (A.4)$$

where  $u = [\text{VEGF}_u]$ ,  $s = [\text{ECM}_s]$  and  $b = [\text{VEGF}_b]$  denote the corresponding concentrations. These equations can be simplified by a number of assumptions that are applicable in the situation under consideration<sup>331</sup>. A first simplifying assumption is that binding sites are part of a large, relatively immobile complex, at least on the time- and length-scale of the FRAP measurement. Then, diffusion of binding sites and bound complexes can be ignored, i.e.,  $D_s = D_b = 0$ . Another simplifying assumption is based on the fact that fluorescence molecules and binding sites are typically at a constant level within the time window in which the FRAP experiment is performed (from seconds to several minutes). Accordingly, we can assume that the system has reached equilibrium before photobleaching. We denote the corresponding equilibrium concentrations of  $\text{VEGF}_u$ ,  $\text{ECM}_s$  and  $\text{VEGF}_b$  by  $u_{eq}$ ,  $s_{eq}$  and  $b_{eq}$ . Although the act of bleaching changes the number of visible free and complexed molecules  $\text{VEGF}_u$  or  $\text{VEGF}_b$ , it does not change the number of free binding sites. Therefore  $s = s_{eq}$  is a constant throughout the photobleaching recovery. This allows us to eliminate equation A.3 and also replace the variable  $s$  in eq. A.2 and A.4 with the constant  $s_{eq}$ . As a result, we can define a pseudo-first-order rate constant as  $k_{on}^* = k_{on} s_{eq}$ . Then, equations A.2, A.3, A.4 reduce to:

$$\frac{\partial u}{\partial t} = D_u \Delta u - k_{on}^* u + k_{off} b \quad (A.5)$$

$$\frac{\partial b}{\partial t} = k_{on}^* u - k_{off} b \quad (A.6)$$

While an analytical solution for A.5 and A.6 can be obtained<sup>331</sup>, here a particular case of that model is used that adequately describes the recovery of the fluorescence in a tiny bleached area near a HUVEC cell. This particular case is commonly known in this context as a reaction-dominant model and describes a scenario where diffusion is very fast compared both to binding and to the timescale of FRAP measurements<sup>331</sup>. In other words, diffusion is not detected in the FRAP recovery, as discussed in Appendix A.2. As a consequence, one can assume that the concentration of free molecules instantly equilibrate after the bleach, i.e.,  $u = u_{eq}$ . Thus, equation A.5 disappears and equation A.6 becomes:

$$\frac{\partial b}{\partial t} = -k_{\text{off}}b + k_{\text{on}}^*U_{\text{eq}} \quad (\text{A.7})$$

Note that the second term on the right hand side in this equation is constant. The solution for this first-order linear equation is given by:

$$b(t) = (k_{\text{on}}^*U_{\text{eq}}/k_{\text{off}}) + Ae^{-k_{\text{off}}t} \quad (\text{A.8})$$

In order to evaluate the constant  $A$  in eq. A.8 recovery data can be normalized to lie between 0 and 1. More precisely, we assume that after normalization the concentration of fluorescence in the bleached zone is zero, i.e.,  $b(0) = 0$ . Then,  $A = -k_{\text{on}}^*U_{\text{eq}}/k_{\text{off}}$ , which leads to:

$$b(t) = k_{\text{on}}^*U_{\text{eq}}/k_{\text{off}} (1 - e^{-k_{\text{off}}t}) \quad (\text{A.9})$$

Due to normalization, we also assume:

$$U_{\text{eq}} + B_{\text{eq}} = 1 \quad (\text{A.10})$$

Equation A.9 yields the behaviour only for the bound complex of the fluorescent protein. Total fluorescence, instead, is given by:

$$\text{frap}(t) = b(t) + u(t) = U_{\text{eq}} + k_{\text{on}}^*U_{\text{eq}}/k_{\text{off}} (1 - e^{-k_{\text{off}}t}) \quad (\text{A.11})$$

Before the bleach, as noted above, the system is at equilibrium, and  $U$  and  $B$  have achieved steady-state values,  $U_{\text{eq}}$  and  $B_{\text{eq}}$ , then:

$$\frac{\partial u}{\partial t} = \frac{\partial b}{\partial t} = 0 \Rightarrow k_{\text{on}}^*U_{\text{eq}} = k_{\text{off}}B_{\text{eq}} \Rightarrow k_{\text{on}}^*U_{\text{eq}}/k_{\text{off}} = B_{\text{eq}} \quad (\text{A.12})$$

Finally, using equations A.10, A.11 and A.12, we obtain:

$$\text{frap}(t) = u(t) + b(t) = 1 - \frac{k_{\text{on}}^*}{k_{\text{on}}^* + k_{\text{off}}} e^{-k_{\text{off}}t} \quad (\text{A.13})$$

This last equation A.13 is the reaction-dominant model used to fit the FRAP recovery data in the ROI in proximity of HUVECs as shown in figure 9.3, b4.

---

## BIBLIOGRAPHY

---

- [1] Bianconi, E. *et al.* An estimation of the number of cells in the human body. *Ann Hum Biol* **40**, 463–471 (2013). p. 1.
- [2] Alberts, B. *et al.* *Molecular Biology of the Cell* (Garland Science, New York, 2002) (Garland, 1997). p. 1.
- [3] Federative Committee on Anatomical Terminology. *Terminologia anatomica* (Thieme, 1998). p. 1.
- [4] Perrimon, N., Pitsouli, C. & Shilo, B.-Z. Signaling mechanisms controlling cell fate and embryonic patterning. *Cold Spring Harb Perspect Biol* **4**, a005975 (2012). p. 1.
- [5] Gumbiner, B. M. Regulation of cadherin-mediated adhesion in morphogenesis. *Nat Rev Mol Cell Biol* **6**, 622–634 (2005). p. 1.
- [6] Halbleib, J. M. & Nelson, W. J. Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. *Genes Dev* **20**, 3199–3214 (2006). p. 1.
- [7] Halder, G., Callaerts, P. & Gehring, W. J. Induction of ectopic eyes by targeted expression of the eyeless gene in drosophila. *Science* **267**, 1788–1792 (1995). p. 1.
- [8] Cohn, M. J., Izpisua-Belmonte, J. C., Abud, H., Heath, J. K. & Tickle, C. Fibroblast growth factors induce additional limb development from the flank of chick embryos. *Cell* **80**, 739–746 (1995). p. 1.
- [9] Sato, T. *et al.* Single Lgr5 stem cells build crypt villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009). p. 2.
- [10] Eiraku, M. *et al.* Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* **472**, 51–56 (2011). p. 2.
- [11] Sasai, Y. Cytosystems dynamics in self-organization of tissue architecture. *Nature* **493**, 318–326 (2013). pp. 2, 4, and 155.
- [12] Salazar-Ciudad, I., Jernvall, J. & Newman, S. A. Mechanisms of pattern formation in development and evolution. *Development* **130**, 2027–2037 (2003). pp. 2 and 155.
- [13] Heisenberg, C.-P. & Solnica-Krezel, L. Back and forth between cell fate specification and movement during vertebrate gastrulation. *Current opinion in genetics & development* **18**, 311–316 (2008). p. 2.
- [14] Uriu, K., Morelli, L. G. & Oates, A. C. Interplay between intercellular signaling and cell movement in development. *Semin Cell Dev Biol* **35C**, 66–72 (2014). pp. 2 and 155.
- [15] Li, D. *et al.* Role of mechanical factors in fate decisions of stem cells. *Regen Med* **6**, 229–240 (2011). p. 2.
- [16] van den Brink, S. C. *et al.* Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* **141**, 4231–4242 (2014). pp. 2 and 150.
- [17] Sampathkumar, A., Yan, A., Krupinski, P. & Meyerowitz, E. M. Physical forces regulate plant development and morphogenesis. *Curr Biol* **24**, R475–R483 (2014). p. 2.

- [18] Lewandoski, M. Conditional control of gene expression in the mouse. *Nat Rev Genet* **2**, 743–755 (2001). p. 2.
- [19] Sander, J. D. & Joung, J. K. Crispr-cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **32**, 347–355 (2014). pp. 2 and 8.
- [20] Carpenter, A. E. & Sabatini, D. M. Systematic genome-wide screens of gene function. *Nat Rev Genet* **5**, 11–22 (2004). p. 2.
- [21] Huisken, J. & Stainier, D. Y. Selective plane illumination microscopy techniques in developmental biology. *Development* **136**, 1963–1975 (2009). pp. 2, 8, and 154.
- [22] Eliceiri, K. W. *et al.* Biological imaging software tools. *Nat Methods* **9**, 697–710 (2012). pp. 2 and 8.
- [23] Ferrell Jr, J. E. Bistability, bifurcations, and Waddington’s epigenetic landscape. *Curr Biol* **22**, R458–R466 (2012). pp. 4, 54, and 56.
- [24] Kondo, S. & Miura, T. Reaction-diffusion model as a framework for understanding biological pattern formation. *Science* **329**, 1616–1620 (2010). pp. 4, 19, and 29.
- [25] Lecuit, T. Developmental mechanics: Cellular patterns controlled by adhesion, cortical tension and cell division. *HFSP Journal* **2**, 72–78 (2008). p. 4.
- [26] Oates, A. C., Gorfinkiel, N., González-Gaitán, M. & Heisenberg, C.-P. Quantitative approaches in developmental biology. *Nat Rev Genet* **10**, 517–530 (2009). pp. 4, 8, and 155.
- [27] Chickarmane, V. *et al.* Computational morphodynamics: a modeling framework to understand plant growth. *Annu Rev Plant Biol* **61**, 65 (2010). pp. 4 and 155.
- [28] Roeder, A. H. *et al.* Computational morphodynamics of plants: integrating development over space and time. *Nat Rev Mol Cell Biol* **12**, 265–273 (2011). pp. 4 and 155.
- [29] Kohl, P. & Noble, D. Systems biology and the virtual physiological human. *Mol Syst Biol* **5** (2009). pp. 5 and 6.
- [30] Bertalanffy, L. v. *General system theory: Foundations, development, applications* (Braziller. New York, 1968). p. 5.
- [31] Kitano, H. Computational systems biology. *Nature* **420**, 206–210 (2002). p. 5.
- [32] Kohl, P., Crampin, E., Quinn, T. & Noble, D. Systems biology: an approach. *Clinical Pharmacology & Therapeutics* **88**, 25–33 (2010). p. 5.
- [33] Kitano, H. Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002). p. 5.
- [34] Westerhoff, H. V. & Palsson, B. O. The evolution of molecular biology into systems biology. *Nat Biotechnol* **22**, 1249–1252 (2004).
- [35] Butcher, E. C., Berg, E. L. & Kunkel, E. J. Systems biology in drug discovery. *Nat Biotechnol* **22**, 1253–1259 (2004).
- [36] Bruggeman, F. J. & Westerhoff, H. V. The nature of systems biology. *Trends Microbiol* **15**, 45–50 (2007). p. 5.
- [37] Noble, D. A theory of biological relativity: no privileged level of causation. *Interface Focus* **2**, 55–64 (2012). p. 6.
- [38] Noble, D. Modeling the heart—from genes to cells to the whole organ. *Science Signaling* **295**, 1678 (2002). p. 6.
- [39] Noble, D. The future: putting Humpty-Dumpty together again. *Biochem Soc Trans* **31**, 156–158 (2003). pp. 6 and 17.

- [40] Noble, D. *The music of life: biology beyond the genome* (OUP Oxford, 2006). p. 6.
- [41] Karr, J. R. *et al.* A whole-cell computational model predicts phenotype from genotype. *Cell* **150**, 389–401 (2012). p. 6.
- [42] Miura, T. Turing and wolpert work together during limb development. *Science signaling* **6**, pe14 (2013). p. 7.
- [43] Bullara, D. & De Decker, Y. Pigment cell movement is not required for generation of turing patterns in zebrafish skin. *Nature communications* **6** (2015). p. 7.
- [44] Turing, A. M. The chemical basis of morphogenesis. 1953. *Bull Math Biol* **52**, 153–97 (1990). pp. 7, 19, 29, 58, and 98.
- [45] Thompson, D. W. *et al.* *On growth and form*. (Cambridge Univ. Press, 1942). p. 7.
- [46] Loeb, J. *Forced movements, tropisms, and animal conduct*, vol. 740 (JB Lippincott, 1918). p. 7.
- [47] Michaelis, L. & Menten, M. Die kinetik der invertinwirkung. *Biochemie* **39**, 333–369 (1913). p. 8.
- [48] Crosetto, N., Bienko, M. & van Oudenaarden, A. Spatially resolved transcriptomics and beyond. *Nat Rev Genet* **16**, 57–66 (2015). p. 8.
- [49] Schwiening, C. J. A brief historical perspective: Hodgkin and Huxley. *J Physiol* **590**, 2571–2575 (2012). p. 14.
- [50] Noble, D. From the Hodgkin-Huxley axon to the virtual heart. *J Physiol* **580**, 15–22 (2007). p. 14.
- [51] Plank, G. *et al.* Generation of histo-anatomically representative models of the individual heart: tools and application. *Philos Trans A Math Phys Eng Sci* **367**, 2257–2292 (2009). p. 14.
- [52] Mirams, G. R., Davies, M. R., Cui, Y., Kohl, P. & Noble, D. Application of cardiac electrophysiology simulations to pro-arrhythmic safety testing. *Br J Pharmacol* **167**, 932–945 (2012). p. 14.
- [53] Wilson, G. Where’s the real bottleneck in scientific computing? *Sci Amer* **94** (2005). p. 14.
- [54] Baxter, S. M., Day, S. W., Fetrow, J. S. & Reisinger, S. J. Scientific software development is not an oxymoron. *PLoS Comput Biol* **2** (2006). p. 14.
- [55] Starrauß, J., de Back, W., Brusch, L. & Deutsch, A. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics* (2014). pp. 15, 23, 25, 26, 68, and 85.
- [56] DiStefano, J. *Dynamic Systems Biology Modeling and Simulation* (Academic Press, 2013). p. 16.
- [57] Eissing, T. *et al.* A computational systems biology software platform for multiscale modeling and simulation: integrating whole-body physiology, disease biology, and molecular reaction networks. *Frontiers in physiology* **2** (2011). p. 16.
- [58] Southern, J. *et al.* Multi-scale computational modelling in biology and physiology. *Prog Biophys Mol Biol* **96**, 60–89 (2008). p. 16.
- [59] Christie, G. R., Nielsen, P. M. F., Blackett, S. A., Bradley, C. P. & Hunter, P. J. Fieldml: concepts and implementation. *Philos. Trans. A Math. Phys. Eng. Sci.* **367**, 1869–1884 (2009). p. 16.
- [60] Bradley, C. *et al.* Opencomiss: a multi-physics & multi-scale computational infrastructure for the vph/physiome project. *Prog Biophys Mol Biol* **107**, 32–47 (2011). p. 16.

- [61] Noble, D. The rise of computational biology. *Nat Rev Mol Cell Biol* **3**, 459–463 (2002). p. 17.
- [62] Walker, D. C. & Southgate, J. The virtual cell: a candidate co-ordinator for "middle-out" modelling of biological systems. *Brief Bioinform* bbp010 (2009). p. 17.
- [63] Hodgkin, A. & Huxley, A. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* **117**, 500–544 (1952). p. 17.
- [64] Dräger, A. *et al.* Modeling metabolic networks in *C. glutamicum*: a comparison of rate laws in combination with various parameter optimization strategies. *BMC Syst Biol* **3**, 5 (2009). p. 19.
- [65] Ferrell, J. E., Jr, Tsai, T. Y.-C. & Yang, Q. Modeling the cell cycle: why do certain circuits oscillate? *Cell* **144**, 874–885 (2011). pp. 19, 28, and 29.
- [66] Cao, J., Qi, X. & Zhao, H. Modeling gene regulation networks using ordinary differential equations. *Methods Mol Biol* **802**, 185–197 (2012). p. 19.
- [67] Hoops, S. *et al.* COPASI—a COMplex Pathway Simulator. *Bioinformatics* **22**, 3067–3074 (2006). p. 19.
- [68] Matsuoka, Y., Funahashi, A., Ghosh, S. & Kitano, H. Modeling and simulation using CellDesigner. *Methods Mol Biol* **1164**, 121–145 (2014). p. 19.
- [69] Wolpert, L. Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* **25**, 1–47 (1969). pp. 19, 28, and 58.
- [70] Gierer, A. & Meinhardt, H. A theory of biological pattern formation. *Kybernetik* **12**, 30–39 (1972). pp. 19, 28, 58, 64, and 98.
- [71] Murray, J. D. How the leopard gets its spots. *Sci Amer* **258**, 80–87 (1988). p. 19.
- [72] Meinhardt, H. Morphogenesis of lines and nets. *Differentiation* **6**, 117–123 (1976). pp. 19 and 98.
- [73] Meinhardt, H. Models of segmentation. In *Somites in developing embryos*, 179–189 (Springer, 1986). p. 19.
- [74] Meinhardt, H. Models for positional signalling with application to the dorsoventral patterning of insects and segregation into different cell types. *Development* **107 Suppl**, 169–180 (1989). pp. 19 and 77.
- [75] Meinhardt, H. Computational modelling of epithelial patterning. *Curr Opin Genet Dev* **17**, 272–280 (2007). p. 19.
- [76] Mori, Y., Jilkine, A. & Edelstein-Keshet, L. Wave-pinning and cell polarity from a bistable reaction-diffusion system. *Biophys J* **94**, 3684–3697 (2008). pp. 19 and 28.
- [77] Anderson, A. & Rejniak, K. (eds.) *Single-Cell-Based Models in Biology and Medicine* (Birkhäuser, 2007). p. 20.
- [78] Deutsch, A. & Dormann, S. *Cellular Automaton Modeling of Biological Pattern Formation: Characterization, Applications, and Analysis (Modeling and Simulation in Science, Engineering and Technology)* (Birkhäuser, 2004). pp. 20 and 21.
- [79] Graner, F. & Glazier, J. A. Simulation of biological cell sorting using a two-dimensional extended potts model. *Phys Rev Lett* **69**, 2013–2016 (1992). pp. 20, 21, 28, 68, and 132.
- [80] Meineke, F., Potten, C. S. & Loeffler, M. Cell migration and organization in the intestinal crypt using a lattice-free model. *Cell Prolif* **34**, 253–266 (2001). pp. 20 and 21.



- [81] Fletcher, A. G., Osterfield, M., Baker, R. E. & Shvartsman, S. Y. Vertex models of epithelial morphogenesis. *Biophys J* **106**, 2291–2304 (2014). pp. 20 and 21.
- [82] Newman, T. J. Modeling multicellular systems using subcellular elements. *Math Biosci Eng* **2**, 613–624 (2005). pp. 20 and 21.
- [83] Drasdo, D. Center-based single-cell models: An approach to multi-cellular organization based on a conceptual analogy to colloidal particles. In Anderson, A., Chaplain, M. & Rejniak, K. (eds.) *Single-Cell-Based Models in Biology and Medicine* (Birkhäuser, Basel, 2007). p. 20.
- [84] Farhadifar, R., Röper, J.-C., Aigouy, B., Eaton, S. & Jülicher, F. The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. *Curr Biol* **17**, 2095–2104 (2007). p. 20.
- [85] Voß-Böhme, A. Multi-scale modeling in morphogenesis: a critical analysis of the cellular Potts model. *PLoS One* **7**, e42852 (2012). pp. 20 and 34.
- [86] Pathmanathan, P. *et al.* A computational study of discrete mechanical tissue models. *Phys Biol* **6**, 036001 (2009). pp. 20 and 22.
- [87] Kang, S., Kahan, S., McDermott, J., Flann, N. & Shmulevich, I. Biocellion: accelerating computer simulation of multicellular biological system models. *Bioinformatics* **30**, 3101–3108 (2014). p. 23.
- [88] Hoehme, S. & Drasdo, D. A cell-based simulation software for multi-cellular systems. *Bioinformatics* **26**, 2641–2642 (2010). pp. 23 and 25.
- [89] Swat, M. H. *et al.* Multi-scale modeling of tissues using compucell3d. *Methods Cell Biol* **110**, 325 (2012). p. 23.
- [90] Mirams, G. R. *et al.* Chaste: an open source C++ library for computational physiology and biology. *PLoS Comput Biol* **9**, e1002970 (2013). p. 23.
- [91] Sütterlin, T., Kolb, C., Dickhaus, H., Jäger, D. & Grabe, N. Bridging the scales: semantic integration of quantitative sbml in graphical multi-cellular models and simulations with episim and copasi. *Bioinformatics* **29**, 223–229 (2013). p. 23.
- [92] Merks, R. M. H., Guravage, M., Inzé, D. & Beemster, G. T. S. Virtualleaf: an open-source framework for cell-based modeling of plant tissue growth and development. *Plant Physiol* **155**, 656–666 (2011). pp. 23 and 25.
- [93] Hucka, M. *et al.* The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531 (2003). pp. 24, 26, 39, and 153.
- [94] de Back, W., Deutsch, A., Drasdo, D., Funahashi, A. & Uhrmacher, A. M. Towards a standard exchange format for spatial, multilevel multicellular models. In Gilbert, D., Heiner, M., Takahashi, K. & Uhrmacher, A. (eds.) *Dagstuhl Reports: Multiscale Spatial Computational Systems Biology*, 214 (Dagstuhl, 2015). pp. 24, 51, and 153.
- [95] Kholodenko, B. N. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur J Biochem* **267**, 1583–1588 (2000). p. 28.
- [96] Glauche, I., Herberg, M. & Roeder, I. Nanog variability and pluripotency regulation of embryonic stem cells—insights from a mathematical model analysis. *PLoS One* **5**, e11238 (2010). p. 28.

- [97] Miyazawa, S., Okamoto, M. & Kondo, S. Blending of animal colour patterns by hybridization. *Nat Commun* **1**, 66 (2010). p. 28.
- [98] Barkley, D. A model for fast computer simulation of waves in excitable media. *Physica D* **49**, 61–70 (1991). p. 28.
- [99] Szabó, A. *et al.* Collective cell motion in endothelial monolayers. *Phys Biol* **7**, 046007 (2010). p. 28.
- [100] Landsberg, K. P. *et al.* Increased cell bond tension governs cell sorting at the drosophila anteroposterior compartment boundary. *Curr Biol* **19**, 1950–1955 (2009). p. 28.
- [101] de Back, W., Zhou, J. X. & Brusch, L. On the role of lateral stabilization during early patterning in the pancreas. *J R Soc Interface* **10**, 20120766 (2013). pp. 28, 87, and 88.
- [102] Köhn-Luque, A. *et al.* Early embryonic vascular patterning by matrix-mediated paracrine signalling: a mathematical model study. *PLoS One* **6**, e24175 (2011). pp. 28, 29, 50, 98, 129, 132, 135, 136, 145, and 146.
- [103] Savill, N. J. & Hogeweg, P. Modelling morphogenesis: from single cells to crawling slugs. *J Theor Biol* **184**, 229–235 (1997). pp. 28 and 33.
- [104] Ma, D. *et al.* Cell packing influences planar cell polarity signaling. *Proc Natl Acad Sci U S A* **105**, 18800–18805 (2008). pp. 28 and 155.
- [105] Meinhardt, H. The algorithmic beauty of sea shells. springer. *Heidelberg, New York* (2003). p. 28.
- [106] Machado, D. *et al.* Modeling formalisms in systems biology. *AMB Express* **1**, 45 (2011). p. 28.
- [107] Wolpert, L. Positional information and pattern formation. *Philos Trans R Soc Lond B Biol Sci* **295**, 441–450 (1981). p. 29.
- [108] Meinhardt, H. & Gierer, A. Applications of a theory of biological pattern formation based on lateral inhibition. *J Cell Sci* **15**, 321–346 (1974). p. 29.
- [109] Balter, A., Merks, R. M., Popławski, N. J., Swat, M. & Glazier, J. A. The glazier-graner-hogeweg model: extensions, future directions, and opportunities for further study. In *Single-Cell-Based Models in Biology and Medicine*, 151–167 (Springer, 2007). p. 31.
- [110] Lecuit, T. & Lenne, P.-F. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nat Rev Mol Cell Biol* **8**, 633–644 (2007). p. 31.
- [111] Glazier, J. A. & Graner, F. Simulation of the differential adhesion driven rearrangement of biological cells. *Phys Rev E* **47**, 2128–2154 (1993). pp. 31 and 109.
- [112] Voß-Böhme, A., Starruß, J. & de Back, W. Cellular Potts model. *Encyclopedia of Systems Biology* 386–390 (2013). p. 31.
- [113] Merks, R. M. H., Perryn, E. D., Shirinifard, A. & Glazier, J. A. Contact-inhibited chemotaxis in de novo and sprouting blood-vessel growth. *PLoS Comput Biol* **4**, e1000163 (2008). pp. 33, 98, 100, 102, 106, 109, 114, 126, and 146.
- [114] Marée, A. F. M., Grieneisen, V. A. & Hogeweg, P. The cellular Potts model and biophysical properties of cells, tissues and morphogenesis. In Anderson, A., Chaplain, M. & Rejniak, K. (eds.) *Single-Cell-Based Models in Biology and Medicine*, 107–136 (Birkhäuser, 2007). p. 33.
- [115] Köhn-Luque, A. *et al.* Dynamics of VEGF matrix-retention in vascular network patterning. *Phys Biol* **10**, 066007 (2013). pp. 34, 50, and 98.

- [116] Chen, N., Glazier, J. A., Izaguirre, J. A. & Alber, M. S. A parallel implementation of the cellular Potts model for simulation of cell-based morphogenesis. *Comput Phys Commun* **176**, 670–681 (2007). p. 36.
- [117] Tapia, J. J. & D'Souza, R. M. Parallelizing the cellular potts model on graphics processing units. *Comput Phys Commun* **182**, 857–865 (2011). p. 36.
- [118] Neal, M. L. *et al.* A reappraisal of how to build modular, reusable models of biological systems. *PLoS Comput Biol* **10**, e1003849 (2014). pp. 41 and 153.
- [119] Herberg, M., Zerjatke, T., de Back, W., Glauch, I. & Roeder, I. Image-based quantification and mathematical modeling of spatial heterogeneity in mesc colonies. *Cytometry Part A* (2015). pp. 50, 150, and 154.
- [120] Waddington, C. H. The strategy of the genes. *London: Allen* **86** (1957). pp. 54 and 56.
- [121] Takahashi, K. & Yamanaka, S. Induced pluripotent stem cells in medicine and biology. *Development* **140**, 2457–2461 (2013). pp. 54 and 55.
- [122] Efrat, S. *Stem cell therapy for Diabetes* (Humana Press, 2010). pp. 54, 62, and 81.
- [123] Ricordi, C., Inverardi, L. & Domínguez-Bendala, J. From cellular therapies to tissue reprogramming and regenerative strategies in the treatment of diabetes. *Regenerative medicine* **7**, 41–48 (2012). p. 54.
- [124] Pagliuca, F. W. & Melton, D. A. How to make a functional  $\beta$ -cell. *Development* **140**, 2472–2483 (2013). p. 54.
- [125] Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D. A. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* **455**, 627–632 (2008). pp. 54, 81, 108, and 124.
- [126] Huang, S., Ernberg, I. & Kauffman, S. Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. In *Seminars in cell & developmental biology*, vol. 20, 869–876 (Elsevier, 2009). p. 55.
- [127] Ben-David, U. & Benvenisty, N. The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nat Rev Cancer* **11**, 268–277 (2011). p. 55.
- [128] O'Doherty, R., Greiser, U. & Wang, W. Nonviral methods for inducing pluripotency to cells. *BioMed res int* **2013** (2013). p. 55.
- [129] Chakraborty, S. *et al.* A CRISPR/Cas9-based system for reprogramming cell lineage specification. *Stem Cell Reports* **3**, 940–947 (2014). p. 55.
- [130] Ke, Q. *et al.* Connexin 43 is involved in the generation of human-induced pluripotent stem cells. *Hum Mol Genet* **22**, 2221–2233 (2013). pp. 55 and 59.
- [131] Redmer, T. *et al.* E-cadherin is crucial for embryonic stem cell pluripotency and can replace oct4 during somatic cell reprogramming. *EMBO Rep* **12**, 720–726 (2011). pp. 55, 59, 80, and 92.
- [132] Davidson, E. H. *et al.* A genomic regulatory network for development. *Science* **295**, 1669–1678 (2002). p. 56.
- [133] Bolouri, H. & Davidson, E. H. Modeling transcriptional regulatory networks. *Bioessays* **24**, 1118–1129 (2002).
- [134] Longabaugh, W. J., Davidson, E. H. & Bolouri, H. Computational representation of developmental genetic regulatory networks. *Dev Biol* **283**, 1–16 (2005).

- [135] Roeder, I. & Radtke, F. Stem cell biology meets systems biology. *Development* **136**, 3525–3530 (2009).
- [136] Enver, T., Pera, M., Peterson, C. & Andrews, P. W. Stem cell states, fates, and the rules of attraction. *Cell Stem Cell* **4**, 387–397 (2009). pp. 64 and 82.
- [137] Macarthur, B. D., Maayan, A. & Lemischka, I. R. Systems biology of stem cell fate and cellular reprogramming. *Nat Rev Mol Cell Biol* **10**, 672–681 (2009). pp. 82 and 93.
- [138] Peltier, J. & Schaffer, D. Systems biology approaches to understanding stem cell fate choice. *Systems Biology, IET* **4**, 1–11 (2010).
- [139] Ladewig, J., Koch, P. & Brüstle, O. Leveling waddington: the emergence of direct programming and the loss of cell fate hierarchies. *Nat Rev Mol Cell Biol* **14**, 225–236 (2013). p. 56.
- [140] Delbrück, M. *Unités Biologiques Douées de Continuité Genetique*, chap. Enzyme systems with alternative steady states, 33–34 (Paris: Editions du CNRS, 1949). p. 56.
- [141] Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**, 318–356 (1961). p. 56.
- [142] Kauffman, S. A. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* **22**, 437–467 (1969). p. 56.
- [143] Huang, S. Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. *Philos Trans R Soc Lond B Biol Sci* **366**, 2247–2259 (2011). p. 56.
- [144] Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000). p. 56.
- [145] Zhou, J. X. & Huang, S. Understanding gene circuits at cell-fate branch points for rational cell reprogramming. *Trends Genet* **27**, 55–62 (2011). pp. 56 and 57.
- [146] Zhou, J. X., Brusch, L. & Huang, S. Predicting pancreas cell fate decisions and reprogramming with a hierarchical multi-attractor model. *PLoS One* **6**, e14752 (2011). pp. 57, 64, 66, 82, and 93.
- [147] Apelqvist, A. *et al.* Notch signalling controls pancreatic cell differentiation. *Nature* **400**, 877–881 (1999). pp. 58, 59, 62, 63, 64, 70, 76, 77, 78, 81, 82, and 91.
- [148] Jensen, J. *et al.* Independent development of pancreatic alpha- and beta-cells from Neurogenin3-expressing precursors: a role for the Notch pathway in repression of premature differentiation. *Diabetes* **49**, 163–176 (2000). pp. 58, 63, 64, 70, 76, 77, 78, and 82.
- [149] Collier, J. R., Monk, N. A., Maini, P. K. & Lewis, J. H. Pattern formation by lateral inhibition with feedback: a mathematical model of delta-notch intercellular signalling. *J Theor Biol* **183**, 429–446 (1996). pp. 58, 63, 64, 70, and 82.
- [150] Monk, N. A. Cell communities and robustness in development. *Bull Math Biol* **59**, 1183–1189 (1997). pp. 64, 73, and 78.
- [151] Lewis, J. Notch signalling and the control of cell fate choices in vertebrates. *Semin Cell Dev Biol* **9**, 583–589 (1998). pp. 58, 63, 70, and 82.
- [152] Baeyens, L. *et al.* In vitro generation of insulin-producing beta cells from adult exocrine pancreatic cells. *Diabetologia* **48**, 49–57 (2005). pp. 59, 63, 81, 82, 89, and 93.
- [153] Baeyens, L. & Bouwens, L. Can beta-cells be derived from exocrine pancreas? *Diabetes Obes Metab* **10 Suppl 4**, 170–178 (2008). pp. 59 and 63.

- [154] Baeyens, L. *et al.* Notch signaling as gatekeeper of rat acinar-to-beta-cell conversion in vitro. *Gastroenterology* **136**, 1750–60.e13 (2009). pp. 59, 63, 77, 81, 88, 91, and 93.
- [155] Marthiens, V., Kazanis, I., Moss, L., Long, K. *et al.* Adhesion molecules in the stem cell niche—more than just staying in shape? *J Cell Sci* **123**, 1613–1622 (2010). p. 59.
- [156] Oyamada, M., Takebe, K., Endo, A., Hara, S. & Oyamada, Y. Connexin expression and gap-junctional intercellular communication in es cells and ips cells. *Frontiers in pharmacology* **4** (2013).
- [157] Pieters, T. & Van Roy, F. Role of cell–cell adhesion complexes in embryonic stem cell biology. *J Cell Sci* **127**, 2603–2613 (2014). p. 59.
- [158] Sprinzak, D. *et al.* Cis-interactions between notch and delta generate mutually exclusive signalling states. *Nature* **465**, 86–90 (2010). p. 61.
- [159] Boareto, M. *et al.* Jagged–delta asymmetry in notch signaling can give rise to a sender/receiver hybrid phenotype. *Proc Natl Acad Sci U S A* 201416287 (2015). pp. 61 and 150.
- [160] Jolly, M. K. *et al.* Operating principles of notch–delta–jagged module of cell–cell communication. *New J Phys* **17**, 055021 (2015). pp. 61 and 150.
- [161] Pin, C. L., Rukstalis, J. M., Johnson, C. & Konieczny, S. F. The bHLH transcription factor *Mist1* is required to maintain exocrine pancreas cell organization and acinar cell identity. *J Cell Biol* **155**, 519–530 (2001). pp. 61 and 79.
- [162] Minami, K., Okano, H., Okumachi, A. & Seino, S. Role of cadherin-mediated cell–cell adhesion in pancreatic exocrine-to-endocrine transdifferentiation. *J Biol Chem* **283**, 13753–13761 (2008). pp. 61, 63, 65, 77, 79, 80, 82, 89, 92, and 93.
- [163] Lammert, E., Cleaver, O. & Melton, D. Induction of pancreatic differentiation by signals from blood vessels. *Science* **294**, 564–567 (2001). pp. 61 and 97.
- [164] Lammert, E., Cleaver, O. & Melton, D. Role of endothelial cells in early pancreas and liver development. *Mech Dev* **120**, 59–64 (2003). p. 61.
- [165] Slack, J. M. Developmental biology of the pancreas. *Development* **121**, 1569–1580 (1995). p. 62.
- [166] Gu, G., Dubauskaite, J. & Melton, D. A. Direct evidence for the pancreatic lineage: *Ngn3*<sup>+</sup> cells are islet progenitors and are distinct from duct progenitors. *Development* **129**, 2447–2457 (2002). pp. 62 and 68.
- [167] Fishman, M. P. & Melton, D. A. Pancreatic lineage analysis using a retroviral vector in embryonic mice demonstrates a common progenitor for endocrine and exocrine cells. *Int J Dev Biol* **46**, 201–207 (2002).
- [168] Zhou, Q. *et al.* A multipotent progenitor domain guides pancreatic organogenesis. *Dev Cell* **13**, 103–114 (2007). p. 62.
- [169] Jensen, J. Gene regulatory factors in pancreatic development. *Dev Dyn* **229**, 176–200 (2004). pp. 62 and 64.
- [170] Gittes, G. K. Developmental biology of the pancreas: a comprehensive review. *Dev Biol* **326**, 4–35 (2009). pp. 62 and 63.
- [171] Habener, J. F., Kemp, D. M. & Thomas, M. K. Minireview: transcriptional regulation in pancreatic development. *Endocrinology* **146**, 1025–1034 (2005). p. 62.
- [172] Schwitzgebel, V. M. *et al.* Expression of *Neurogenin3* reveals an islet cell precursor population in the pancreas. *Development* **127**, 3533–3542 (2000). pp. 62, 63, and 76.

- [173] Gradwohl, G., Dierich, A., LeMeur, M. & Guillemot, F. Neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc Natl Acad Sci U S A* **97**, 1607–1611 (2000). pp. 62, 63, 76, 82, and 83.
- [174] Krapp, A. *et al.* The p48 DNA-binding subunit of transcription factor PTF1 is a new exocrine pancreas-specific basic helix-loop-helix protein. *EMBO J* **15**, 4317–4329 (1996). pp. 63, 64, and 82.
- [175] Kawaguchi, Y. *et al.* The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet* **32**, 128–134 (2002). pp. 66, 82, and 88.
- [176] Zecchin, E. *et al.* Evolutionary conserved role of Ptf1a in the specification of exocrine pancreatic fates. *Dev Biol* **268**, 174–184 (2004). pp. 63, 64, and 82.
- [177] Jørgensen, M. C. *et al.* An illustrated review of early pancreas development in the mouse. *Endocr Rev* **28**, 685–705 (2007). pp. 63, 76, and 88.
- [178] Kesavan, G. *et al.* Cdc42-mediated tubulogenesis controls cell specification. *Cell* **139**, 791–801 (2009). p. 63.
- [179] Kim, S. K. & Hebrok, M. Intercellular signals regulating pancreas development and function. *Genes Dev* **15**, 111–127 (2001). pp. 63 and 79.
- [180] Lammert, E., Brown, J. & Melton, D. A. Notch gene expression during pancreatic organogenesis. *Mech Dev* **94**, 199–203 (2000).
- [181] Fujikura, J. *et al.* Notch/Rbp-j signaling prevents premature endocrine and ductal cell differentiation in the pancreas. *Cell Metab* **3**, 59–65 (2006). p. 64.
- [182] Kim, W., Shin, Y.-K., Kim, B.-J. & Egan, J. M. Notch signaling in pancreatic endocrine cell and diabetes. *Biochem Biophys Res Commun* **392**, 247–251 (2010). p. 63.
- [183] Henrique, D. *et al.* Maintenance of neuroepithelial progenitor cells by delta-notch signalling in the embryonic chick retina. *Curr Biol* **7**, 661–670 (1997). p. 63.
- [184] Cohen, D. E. & Melton, D. Turning straw into gold: directing cell fate for regenerative medicine. *Nat Rev Genet* **12**, 243–252 (2011). p. 63.
- [185] Hall, P. A. & Lemoine, N. R. Rapid acinar to ductal transdifferentiation in cultured human exocrine pancreas. *J Pathol* **166**, 97–103 (1992). p. 63.
- [186] Pinho, A. V. *et al.* Adult pancreatic acinar cells dedifferentiate to an embryonic progenitor phenotype with concomitant activation of a senescence programme that is present in chronic pancreatitis. *Gut* **60**, 958–966 (2011). p. 63.
- [187] Baeyens, L., Rومان, I. & Bouwens, L. *Stem cell therapy for diabetes*, chap. Generation of beta cells from acinar cells, 153–166 (Humana Press, 2010). pp. 63, 81, 82, and 88.
- [188] Zhu, L. *et al.* Inhibition of Mist1 homodimer formation induces pancreatic acinar-to-ductal metaplasia. *Mol Cell Biol* **24**, 2673–2681 (2004). pp. 63, 65, 77, 79, 80, 82, and 92.
- [189] Huang, S., Guo, Y.-P., May, G. & Enver, T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev Biol* **305**, 695–713 (2007). pp. 64, 82, and 93.
- [190] Jacquemin, P. *et al.* Transcription factor hepatocyte nuclear factor 6 regulates pancreatic endocrine cell differentiation and controls expression of the proendocrine gene *ngn3*. *Mol Cell Biol* **20**, 4445–4454 (2000). pp. 64 and 83.
- [191] Haumaitre, C. *et al.* Lack of TCF2/vHNF1 in mice leads to pancreas agenesis. *Proc Natl Acad Sci U S A* **102**, 1490–1495 (2005). pp. 64 and 83.

- [192] Maestro, M. A. *et al.* Hnf6 and Tcf2 (MODY5) are linked in a gene network operating in a precursor cell domain of the embryonic pancreas. *Hum Mol Genet* **12**, 3307–3314 (2003). pp. 64 and 83.
- [193] Lee, J. C. *et al.* Regulation of the pancreatic pro-endocrine gene neurogenin3. *Diabetes* **50**, 928–936 (2001). p. 64.
- [194] Chiang, M.-K. & Melton, D. A. Single-cell transcript analysis of pancreas development. *Dev Cell* **4**, 383–393 (2003). pp. 65, 83, and 88.
- [195] Sander, M. *et al.* Homeobox gene Nkx6.1 lies downstream of Nkx2.2 in the major pathway of beta-cell formation in the pancreas. *Development* **127**, 5533–5540 (2000). pp. 65 and 83.
- [196] Wilson, M. E., Scheel, D. & German, M. S. Gene expression cascades in pancreatic development. *Mech Dev* **120**, 65–80 (2003). pp. 65 and 83.
- [197] Schaffer, A. E., Freude, K. K., Nelson, S. B. & Sander, M. Nkx6 transcription factors and ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. *Dev Cell* **18**, 1022–1029 (2010). pp. 65 and 83.
- [198] Angeli, D., Ferrell Jr., J. E. & Sontag, E. D. Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc Natl Acad Sci U S A* **101**, 1822–1827 (2004). p. 66.
- [199] Gittes, G. K. & Rutter, W. J. Onset of cell-specific gene expression in the developing mouse pancreas. *Proc Natl Acad Sci U S A* **89**, 1128–1132 (1992). p. 68.
- [200] de Boer, R. & Pagie, L. *GRIND: GReat INtegrator Differential equations*. Theoretical Biology and Bioinformatics, Utrecht University (2011). pp. 68 and 85.
- [201] Ermentrout, B. *Simulating, analyzing, and animating dynamical systems: a guide to XPPAUT for researchers and students*. Software, Environments, and Tools (Society for Industrial and Applied Mathematics (SIAM), 2002). pp. 68 and 85.
- [202] Haddon, C., Jiang, Y. J., Smithers, L. & Lewis, J. Delta-notch signalling and the patterning of sensory cell differentiation in the zebrafish ear: evidence from the mind bomb mutant. *Development* **125**, 4637–4644 (1998). p. 70.
- [203] Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770–776 (1999). p. 70.
- [204] Masui, T. *et al.* Transcriptional autoregulation controls pancreatic Ptf1a expression during development and adulthood. *Mol Cell Biol* **28**, 5458–5468 (2008). p. 72.
- [205] Cohen, M., Georgiou, M., Stevenson, N. L., Miodownik, M. & Baum, B. Dynamic filopodia transmit intermittent Delta-Notch signaling to drive pattern refinement during lateral inhibition. *Dev Cell* **19**, 78–89 (2010). pp. 75 and 76.
- [206] Murtaugh, L., Law, A., Dor, Y. & Melton, D. A.  $\beta$ -catenin is essential for pancreatic acinar but not islet development. *Development* **132**, 4663–4674 (2005). p. 77.
- [207] Wells, J. M. *et al.* Wnt/beta-catenin signaling is required for development of the exocrine pancreas. *BMC Dev Biol* **7**, 4 (2007). p. 77.
- [208] Heiser, P. W., Lau, J., Taketo, M. M., Herrera, P. L. & Hebrok, M. Stabilization of beta-catenin impacts pancreas growth. *Development* **133**, 2023–2032 (2006). p. 77.
- [209] Hart, A., Papadopoulou, S. & Edlund, H. Fgf10 maintains notch activation, stimulates proliferation, and blocks differentiation of pancreatic epithelial cells. *Dev Dyn* **228**, 185–193 (2003). p. 77.

- [210] Norgaard, G. A., Jensen, J. N. & Jensen, J. Fgf10 signaling maintains the pancreatic progenitor cell state revealing a novel role of Notch in organ development. *Dev Biol* **264**, 323–338 (2003). pp. 77 and 88.
- [211] Murtaugh, L. C., Stanger, B. Z., Kwan, K. M. & Melton, D. A. Notch signaling controls multiple steps of pancreatic differentiation. *Proc Natl Acad Sci U S A* **100**, 14920–14925 (2003). pp. 77 and 78.
- [212] Gurdon, J. B., Lemaire, P. & Kato, K. Community effects and related phenomena in development. *Cell* **75**, 831–834 (1993). p. 78.
- [213] Saka, Y., Lhoussaine, C., Kuttler, C., Ullner, E. & Thiel, M. Theoretical basis of the community effect in development. *BMC Syst Biol* **5**, 54 (2011). p. 78.
- [214] Bolouri, H. & Davidson, E. H. The gene regulatory network basis of the community effect and analysis of a sea urchin embryo example. *Dev Biol* **340**, 170–178 (2010). p. 78.
- [215] Mangold, O. & Spemann, H. Über induktion von medullarplatte durch medullarplatte im jüngeren keim, ein beispiel homöogenetischer oder assimilatorischer induktion. *Wilhelm Roux Arch. EntwMech. Org.* **111**, 341–422 (1927). pp. 78 and 88.
- [216] Hald, J. *et al.* Activated Notch1 prevents differentiation of pancreatic acinar cells and attenuate endocrine development. *Dev Biol* **260**, 426–437 (2003). p. 78.
- [217] Afelik, S. *et al.* Notch-mediated patterning and cell fate allocation of pancreatic progenitor cells. *Development* **139**, 1744–1753 (2012).
- [218] Ahnfelt-Rønne, J. *et al.* Ptf1a-mediated control of dll1 reveals an alternative to the lateral inhibition mechanism. *Development* **139**, 33–45 (2012). p. 78.
- [219] Stadtfeld, M. & Hochedlinger, K. Induced pluripotency: history, mechanisms, and applications. *Genes Dev* **24**, 2239–2263 (2010). p. 80.
- [220] Plath, K. & Lowry, W. E. Progress in understanding reprogramming to the induced pluripotent state. *Nat Rev Genet* **12**, 253–265 (2011). p. 80.
- [221] Cherry, A. B. C. & Daley, G. Q. Reprogramming cellular identity for regenerative medicine. *Cell* **148**, 1110–1122 (2012). p. 80.
- [222] Sancho-Martinez, I., Baek, S. H. & Izpisua Belmonte, J. C. Lineage conversion methodologies meet the reprogramming toolbox. *Nat Cell Biol* **14**, 892–899 (2012). p. 80.
- [223] Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006). p. 80.
- [224] Cohen, M., Baum, B. & Miodownik, M. The importance of structured noise in the generation of self-organizing tissue patterns through contact-mediated cell-cell signalling. *J R Soc Interface* **8**, 787–798 (2011). pp. 80 and 92.
- [225] Lin, J. *et al.* Microenvironment-evoked cell lineage conversion: Shifting the focus from internal reprogramming to external forcing. *Ageing Res Rev* (2012). pp. 80 and 92.
- [226] Li, L., Bennett, S. A. L. & Wang, L. Role of e-cadherin and other cell adhesion molecules in survival and differentiation of human pluripotent stem cells. *Cell Adh Migr* **6**, 59–70 (2012). pp. 80 and 92.
- [227] Chen, T. *et al.* E-cadherin-mediated cell-cell contact is critical for induced pluripotent stem cell generation. *Stem Cells* **28**, 1315–1325 (2010). pp. 108, 109, 114, 120, 124, 128, 129, 133, and 146.



- [228] Zhang, J. *et al.* Cortical neural precursors inhibit their own differentiation via n-cadherin maintenance of beta-catenin signaling. *Dev Cell* **18**, 472–479 (2010).
- [229] Soncin, F. *et al.* Abrogation of e-cadherin-mediated cell-cell contact in mouse embryonic stem cells results in reversible lif-independent self-renewal. *Stem Cells* **27**, 2069–2080 (2009). pp. 80 and 92.
- [230] Rooman, I., Heremans, Y., Heimberg, H. & Bouwens, L. Modulation of rat pancreatic acinoductal transdifferentiation and expression of pdx-1 in vitro. *Diabetologia* **43**, 907–914 (2000). pp. 81, 82, and 93.
- [231] Song, K.-H. *et al.* In vitro transdifferentiation of adult pancreatic acinar cells into insulin-expressing cells. *Biochem Biophys Res Commun* **316**, 1094–1100 (2004). pp. 81 and 93.
- [232] Minami, K. *et al.* Lineage tracing and characterization of insulin-secreting cells generated from adult pancreatic acinar cells. *Proc Natl Acad Sci U S A* **102**, 15116–15121 (2005). pp. 81, 82, 88, and 89.
- [233] Baeyens, L. *et al.* Ngn3 expression during postnatal in vitro beta cell neogenesis induced by the JAK/STAT pathway. *Cell Death Differ* **13**, 1892–1899 (2006). pp. 81 and 89.
- [234] Schittler, D., Hasenauer, J., Allgöwer, F. & Waldherr, S. Cell differentiation modeled via a coupled two-switch regulatory network. *Chaos* **20**, 045121 (2010). p. 82.
- [235] Glauche, I., Cross, M., Loeffler, M. & Roeder, I. Lineage specification of hematopoietic stem cells: mathematical modeling and biological implications. *Stem Cells* **25**, 1791–1799 (2007). p. 82.
- [236] Ahlgren, U., Pfaff, S. L., Jessell, T. M., Edlund, T. & Edlund, H. Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature* **385**, 257–260 (1997). p. 82.
- [237] Monahan-Earley, R., Dvorak, A. & Aird, W. Evolutionary origins of the blood vascular system and endothelium. *J Thromb Haemost* **11**, 46–66 (2013). p. 96.
- [238] Aird, W. C. Spatial and temporal dynamics of the endothelium. *J Thromb Haemost* **3**, 1392–1406 (2005). p. 96.
- [239] Gonzalez-Crussi, F. Vasculogenesis in the chick embryo. an ultrastructural study. *Am J Anat* **130**, 441–459 (1971). p. 97.
- [240] Matsumoto, K., Yoshitomi, H., Rossant, J. & Zaret, K. S. Liver organogenesis promoted by endothelial cells prior to vascular function. *Science* **294**, 559–563 (2001). p. 97.
- [241] Herrero, M. A., Köhn, A. & Pérez-Pomares, J. M. Modelling vascular morphogenesis: Current views on blood vessels development. *Math Mod Meth Appl Sci* **19**, Suppl, 1483–1537 (2009). pp. 97, 98, 105, and 145.
- [242] Czirok, A. & Little, C. D. Pattern formation during vasculogenesis. *Birth Defects Research Part C: Embryo Today: Reviews* **96**, 153–162 (2012). pp. 97 and 151.
- [243] Ambrosi, D., Bussolino, F. & Preziosi, L. A review of vasculogenesis models. *J. Theor. Med.* **6**, 1–19 (2005). p. 105.
- [244] Scianna, M., Bell, C. & Preziosi, L. A review of mathematical models for the formation of vascular networks. *J Theor Biol* **333**, 174–209 (2013). pp. 98, 145, and 151.
- [245] Koch, A. & Meinhardt, H. Biological pattern formation: from basic mechanisms to complex structures. *Reviews of modern physics* **66**, 1481 (1994). pp. 98 and 100.

- [246] Murray, J. D., Oster, G. F. & Harris, A. K. A mechanical model for mesenchymal morphogenesis. *J Math Biol* **17**, 125–129 (1983). p. 98.
- [247] Murray, J., Manoussaki, D., Lubkin, S. & Vernon, R. A mechanical theory of in vitro vascular network formation. In *Vascular morphogenesis: in vivo, in vitro, in mente*, 173–188 (Springer, 1996).
- [248] Manoussaki, D., Lubkin, S. R., Vernon, R. B. & Murray, J. D. A mechanical model for the formation of vascular networks *in vitro*. *Acta Biotheor* **44**, 271–282 (1996). pp. 100, 109, and 145.
- [249] van Oers, R. F., Rens, E. G., LaValley, D. J., Reinhart-King, C. A. & Merks, R. M. Mechanical cell-matrix feedback explains pairwise and collective endothelial cell behavior in vitro. *PLoS Comput Biol* **10**, e1003774 (2014). pp. 98, 100, and 151.
- [250] Serini, G. *et al.* Modeling the early stages of vascular network assembly. *The EMBO Journal* **22**, 1771–1779 (2003). pp. 98, 100, 106, 145, and 146.
- [251] Gamba, A. *et al.* Percolation, morphogenesis, and Burgers dynamics in blood vessel formation. *Phys Rev Lett* **90**, 118101 (2003). pp. 100, 109, 126, 145, and 146.
- [252] Merks, R. & Glazier, J. Dynamic mechanisms of blood vessel growth. *Nonlinearity* **19**, C1–C10 (2006). pp. 100, 106, 109, 113, 114, 118, 121, 126, 132, 135, and 136.
- [253] Merks, R. M. H., Brodsky, S. V., Goligorsky, M. S., Newman, S. A. & Glazier, J. A. Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodelling. *Dev Biol* **289**, 44–54 (2006). pp. 98, 100, 102, 145, and 146.
- [254] Merks, R. M. H. & Glazier, J. A. Dynamics mechanisms of blood vessel growth. *Nonlinearity* **19**, C1–C10 (2006). pp. 100, 102, and 114.
- [255] Merks, R. M. H., Newman, S. A. & Glazier, J. A. Cell oriented modelling of in vitro capillary development. *LNCS* **3305**, 425–434 (2004). pp. 100, 106, and 114.
- [256] Risau, W. & Flamme, I. Vasculogenesis. *Annu Rev Cell Dev Biol* **11**, 73–91 (1995). pp. 101, 105, 106, 108, 124, 126, and 128.
- [257] Nico, B., Vacca, A., Giorgis, M. D., Roncali, L. & Ribatti, D. Vascular endothelial growth factor and vascular endothelial growth factor receptor-2 expression in the chick embryo area vasculosa. *Histochem J* **33**, 283–286 (2001). pp. 101, 106, 107, 108, 126, and 146.
- [258] Arnaoutova, I. & Kleinman, H. K. In vitro angiogenesis: endothelial cell tube formation on gelled basement membrane extract. *Nature protocols* **5**, 628–635 (2010). p. 101.
- [259] Georges, A. L. & Nugent, M. A. pH regulates vascular endothelial growth factor binding to fibronectin: A mechanism for control of extracellular matrix storage and release. *J Biol Chem* **279**, 2307–2315 (2004). pp. 101 and 108.
- [260] Davis, G. E. & Senger, D. R. Endothelial extracellular matrix: Biosynthesis, remodelling and functions during vascular morphogenesis and neovessel stabilization. *Circ Res* **97**, 1093–1107 (2005).
- [261] Stringer, S. The role of heparan sulphate proteoglycans in angiogenesis. *Biochem Soc Trans* **34**, 451–453 (2006). pp. 101, 106, 107, 108, 124, 126, and 133.
- [262] Shamloo, A., Ma, N., Poo, M., Sohn, L. & Heilshorn, S. Endothelial cell polarization and chemotaxis in a microfluidic device. *Lab Chip* **8**, 1292–1299 (2008). pp. 103, 135, 136, and 138.

- [263] Szabó, A., Perryn, E. D. & Czirók, A. Network formation of tissue cells via preferential attraction to elongated structures. *Phys Rev Lett* **98** (2007). 038102. pp. 103, 109, and 147.
- [264] Palm, M. M. & Merks, R. M. Vascular networks due to dynamically arrested crystalline ordering of elongated cells. *Physical Review E* **87**, 012725 (2013). pp. 103 and 151.
- [265] Daramola, O., Heyderman, R., Klein, N., Shennan, G. & Levin, M. Detection of fibronectin expression by human endothelial cells using an enzyme-linked immunosorbent assay (ELISA): enzymatic degradation by activated plasminogen. *J Immunol Methods* **202**, 67–75 (1997). p. 104.
- [266] Bayless, K. J., Salazar, R. & Davis, G. E. Rgd-dependent vacuolation and lumen formation observed during endothelial cell morphogenesis in three-dimensional fibrin matrices involves the  $\alpha_v\beta_3$  and  $\alpha_5\beta_1$  integrins. *The American journal of pathology* **156**, 1673–1683 (2000). pp. 104 and 151.
- [267] Califano, J. P. & Reinhart-King, C. A. Substrate stiffness and cell area predict cellular traction stresses in single cells and cells in contact. *Cell. Mol. Bioeng.* **3**, 68–75 (2010). p. 104.
- [268] Sato, Y. *et al.* Dynamic analysis of vascular morphogenesis using transgenic quail embryos. *PloS one* **5**, e12674 (2010). p. 104.
- [269] Jakobsson, L. *et al.* Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nat Cell Biol* **12**, 943–953 (2010). p. 147.
- [270] Arima, S. *et al.* Angiogenic morphogenesis driven by dynamic and heterogeneous collective endothelial cell movement. *Development* **138**, 4763–4776 (2011). p. 104.
- [271] Cleaver, O. & Krieg, P. Vascular development. In Rosenthal, N. & Harvey, R. (eds.) *Heart development and regeneration*, 487–528 (Academic Press, San Diego, 2010). p. 105.
- [272] Drake, D. J., Brandt, S. J., Trusk, T. C. & Little, C. D. TAL1/SCL is expressed in endothelial progenitor cells/angioblasts and defines a dorsal-to-ventral gradient of vasculogenesis. *Dev Biol* **192**, 17–30 (1997). p. 105.
- [273] Wolpert, L. *et al.* *Principles of Development* (Oxford University Press, Oxford, 2006), third edn. p. 106.
- [274] Yamaguchi, T. P., Dumont, D. J., Conlon, R. A., Breitman, M. L. & Rossant, J. Flk-1 and Flt-related receptor tyrosine kinase is an early marker for endothelial cell precursors. *Development* **118**, 489–498 (1993). pp. 106 and 107.
- [275] Volin, M. V., Joseph, L., Shockley, M. S. & Davies, P. F. Chemokine receptor CXCR4 expression in endothelium. *Biochem Biophys Res Co* **242**, 46–53 (1998). p. 106.
- [276] Dumont, D. J. *et al.* Vascularization of the mouse embryo: a study of Flk-1, Tek, Tie, and vascular endothelial growth factor expression during development. *Dev Dyn* **203**, 80–92 (1995). pp. 106 and 124.
- [277] Poole, T., Finkelstein, E. & Cox, C. The role of FGF and VEGF in angioblast induction and migration during vascular development. *Dev Dyn* **220**, 1–17 (2001). pp. 106, 107, and 128.
- [278] Miura, T. & Tanaka, R. *In Vitro* vasculogenesis models revisited - measurement of VEGF diffusion in matrigel. *Math Model Nat Phenom* **4**, 118–130 (2009). pp. 106, 114, and 141.
- [279] Risau, W. & Lemmon, V. Changes in the vascular extracellular matrix during embryonic vasculogenesis and angiogenesis. *Dev Biol* **125**, 441–450 (1988). pp. 106, 108, and 124.

- [280] Larsson, J. *et al.* Abnormal angiogenesis but intact hematopoietic potential in TGF- $\beta$  type I receptor-deficient mice. *EMBO J* **20**, 1663–1673 (2001). p. 108.
- [281] Wijelath, E. *et al.* Heparin-ii domain of fibronectin is a vascular endothelial growth factor-binding domain: enhancement of VEGF biological activity by a singular growth factor/matrix protein synergism. *Circ Res* **99**, 853–860 (2006). pp. 106, 108, 124, 128, and 133.
- [282] Watt, S. M., Gschmeissner, S. E. & Bates, P. A. PECAM-1: Its expression and function as a cell adhesion molecule on hemopoietic and endothelial cells. *Leuk Lymphoma* **17**, 229–244 (1995). p. 107.
- [283] Davis, S. *et al.* Isolation of angiopoietin-1, a ligand for the TIE2 receptor, by secretion-trap expression cloning. *Cell* **87**, 1161–1169 (1996). p. 107.
- [284] Houck, K., Leung, D., Rowland, A., Winer, J. & Ferrara, N. Dual regulation of vascular endothelial growth factor bioavailability by genetic and proteolytic mechanisms. *J Biol Chem* **267**, 26031–26037 (1992). pp. 107, 109, and 133.
- [285] Park, J. E., Keller, G. A. & Ferrara, N. The vascular endothelial growth factor (VEGF) isoforms: Differential deposition into the subepithelial extracellular matrix and bioactivity of extracellular matrix-bound VEGF. *Mol Biol Cell* **4**, 1317–1326 (1993).
- [286] Ferrara, N. Molecular and biological properties of vascular endothelial growth factor. *J Mol Med* **77**, 527–543 (1999). pp. 107 and 126.
- [287] Drake, C. J., LaRue, A., Ferrara, N. & Little, C. D. VEGF regulates cell behavior during vasculogenesis. *Dev Biol* **224**, 178–188 (2000). pp. 107 and 126.
- [288] Vokes, S. A. & Krieg, P. A. Endoderm is required for vascular endothelial tube formation, but not for angioblast specification. *Development* **129**, 775–785 (2002). p. 107.
- [289] Flamme, I., Breier, G. & Risau, W. Vascular endothelial growth factor (VEGF) and VEGF receptor 2 (Flk-1) are expressed during vasculogenesis and vascular differentiation in the quail embryo. *Dev Biol* **169**, 699–712 (1995). p. 107.
- [290] Neufeld, G., Cohen, T., Gengrinovitch, S. & Poltorak, Z. Vascular endothelial growth factor (VEGF) and its receptors. *FASEB J* **13**, 9–22 (1999). p. 107.
- [291] Wu, L. W. *et al.* Utilization of distinct signalling pathways by receptors for vascular endothelial cell growth factor and other mitogens in the induction of endothelial cell proliferation. *J Biol Chem* **275**, 5096–5103 (2000). p. 108.
- [292] Ruhrberg, C. *et al.* Spatially restricted patterning cues provided by heparin-binding VEGF-A control blood vessel branching morphogenesis. *Genes Dev* **16**, 2684–2698 (2002). pp. 108, 109, 114, 120, 124, 126, 128, 129, 133, and 146.
- [293] Zamir, E. A., Czirók, A., Cui, C., Little, C. D. & Rongish, B. J. Mesodermal cell displacements during avian gastrulation are due to both individual cell-autonomous and convective tissue movements. *P Natl Acad Sci USA* **103**, 9806–1981 (2006). 11. p. 108.
- [294] Argraves, W. S. & Drake, C. J. Genes critical to vasculogenesis as defined by systematic analysis of vascular defects in knockout mice. *Anat Rec Part A* **286A**, 875–884 (2005). p. 108.
- [295] Ashikari-Hada, S., Habuchi, H., Kariya, Y. & Kimata, K. Heparin regulates vascular endothelial growth factor 165-dependent mitogenic activity, tube formation, and its receptor phosphorylation of human endothelial cells. *J Biol Chem* **35**, 1508–3151 (2005). Comparison of the effects of heparin and modified heparins. p. 108.

- [296] Soker, S., Svahn, C. & Neufeld, G. Vascular endothelial growth factor is inactivated by binding to alpha 2-macroglobulin and the binding is inhibited by heparin. *J Biol Chem* **268**, 7685–7691 (1993). pp. 109 and 133.
- [297] Breier, G., Albrecht, U., Sterrer, S. & Risau, W. Expression of vascular endothelial growth factor during embryonic angiogenesis and endothelial cell differentiation. *Development* **114**, 521–532 (1992). pp. 109 and 146.
- [298] Millauer, B. *et al.* High affinity VEGF binding and developmental expression suggest Flk-1 as a major regulator of vasculogenesis and angiogenesis. *Cell* **72**, 835–846 (1993).
- [299] Lai, L., Bohnsack, L., Niederreither, K. & Hirschi, K. K. Retinoic acid regulates endothelial cell proliferation during vasculogenesis. *Development* **120**, 6465–6474 (2003). p. 109.
- [300] Drake, C. J. & Little, C. D. Exogenous vascular endothelial growth factor induces malformed and hyperfused vessels during embryonic neovascularization. *Proc Natl Acad Sci U S A* **92**, 7657–7661 (1995). p. 109.
- [301] Meinhardt, H. Biological pattern formation as a complex dynamic phenomenon. *Int J Bif Chaos* **7**, 1–26 (1997). p. 109.
- [302] Szabó, A., Mehes, E., Kosa, E. & Czirók, A. Multicellular sprouting *in vitro*. *Biophys J* **95**, 2702–2710 (2008). pp. 109 and 147.
- [303] Beysens, D. A., Forgacs, G. & Glazier, J. A. Cell sorting is analogous to phase ordering in fluids. *Proc Natl Acad Sci USA* **97**, 9467–9471 (2000). p. 111.
- [304] Robins, V. Computational topology for point data: Betti numbers of  $\alpha$ -shapes. *Lect Notes Phys* 261–274 (2002). p. 113.
- [305] Guidolin, D. *et al.* Order and disorder in the vascular network. *Leukemia* **18**, 1745–1750 (2004). p. 113.
- [306] Mancardi, D., Varetto, G., Bucci, E., Maniero, F. & Guiot, C. Fractal parameters and vascular networks: facts & artifacts. *Theor Biol Med Model* **5**, 12 (2008). p. 113.
- [307] Kirchner, L. & Gruber, P. S. B. Quantitation of angiogenesis in the chick chorioallantoic membrane model using fractal analysis. *Microvasc Res* **51**, 2–14 (1996). p. 113.
- [308] Smith, T. G., Lange, G. D. & Marks, W. B. Fractal methods and results in cellular morphology - dimensions, lacunarity and multifractals. *J Neurosci Meth* **69**, 123–136 (1996). p. 113.
- [309] Tolle, C., McJunkin, T. & Gorsich, D. An efficient implementation of the gliding box lacunarity algorithm. *Physica D* **237**, 306–315 (2008). p. 113.
- [310] Karperien, A. Fraclac for ImageJ, version 2 (1999–2007). p. 113.
- [311] Tosin, A., Ambrosi, D. & Preziosi, L. Mechanics and chemotaxis in the morphogenesis of vascular networks. *Bull Math Biol* **68**, 1819–1836 (2006). pp. 121, 126, and 145.
- [312] Zajac, M., Jones, G. L. & Glazier, J. A. Simulation of convergent extension by way of anisotropic differential adhesion. *J Theor Biol* **222**, 247–259 (2003). p. 121.
- [313] Francis, S. E. *et al.* Central roles of  $\alpha_5\beta_1$  integrin and fibronectin in vascular development in mouse embryos and embryoid bodies. *Arterioscler Thromb Vasc Biol* **22**, 927–933 (2002). p. 126.
- [314] Carmeliet, P. Angiogenesis in life, disease and medicine. *Nature* **438**, 932–936 (2005). p. 128.

- [315] Novosel, E., Kleinhans, C. & Kluger, P. Vascularization is the key challenge in tissue engineering. *Adv Drug Deliv Rev* **63**, 300–311 (2011). pp. 128 and 151.
- [316] Ferrara, N., Gerber, H. & LeCouter, J. The biology of vegf and its receptors. *Nat Med* **9**, 669–676 (2003). p. 128.
- [317] Olsson, A., Dimberg, A., Kreuger, J. & Claesson-Welsh, L. VEGF receptor signalling - in control of vascular function. *Nat Rev Mol Cell Biol* **7**, 359–371 (2006). p. 128.
- [318] Herbert, S. & Stainier, D. Molecular control of endothelial cell behaviour during blood vessel morphogenesis. *Nat Rev Mol Cell Biol* **12**, 551–564 (2011). p. 128.
- [319] Cleaver, O. & Krieg, P. VEGF mediates angioblast migration during development of the dorsal aorta in xenopus. *Development* **125**, 3905–3914 (1998). p. 128.
- [320] Gerhardt, H. *et al.* Vegf guides angiogenic sprouting utilizing endothelial tip cell filopodia. *J Cell Biol* **161**, 1163–1177 (2003).
- [321] Hogan, K., Ambler, C., Chapman, D. & Bautch, V. The neural tube patterns vessels developmentally using the VEGF signaling pathway. *Development* **131**, 1503–1513 (2004). p. 128.
- [322] Stalmans, I. *et al.* Arteriolar and venular patterning in retinas of mice selectively expressing vegf isoforms. *J Clin Invest* **109**, 327–336 (2002). p. 128.
- [323] Lee, S., Jilani, S., Nikolova, G., Carpizo, D. & Iruela-Arispe, M. Processing of vegf-a by matrix metalloproteinases regulates bioavailability and vascular patterning in tumors. *J Cell Biol* **169**, 681–691 (2005).
- [324] Stenzel, D. *et al.* Integrin-dependent and -independent functions of astrocytic fibronectin in retinal angiogenesis. *Development* **138**, 4451–4463 (2011). pp. 128 and 129.
- [325] Clark, R., Folkvord, J. & Nielsen, L. Either exogenous or endogenous fibronectin can promote adherence of human endothelial cells. *J. Cell. Sci.* **82**, 263–280 (1986). p. 133.
- [326] Reinhart-King, C. A., Dembo, M. & Hammer, D. A. The dynamics and mechanics of endothelial cell spreading. *Biophys J* **89**, 676–689 (2005).
- [327] Mitsi, M., Hong, Z., Costello, C. & Nugent, M. Heparin-mediated conformational changes in fibronectin expose vascular endothelial growth factor binding sites. *Biochemistry (Mosc)* **45**, 10319–10328 (2006). pp. 137 and 143.
- [328] Zhou, X. *et al.* Fibronectin fibrillogenesis regulates three-dimensional neovessel formation. *Genes Dev* **22**, 1231–1243 (2008). p. 133.
- [329] Peters, R., Peters, J., Tews, K. & Bähr, W. A microfluorimetric study of translational diffusion in erythrocyte membranes. *Biochim Biophys Acta* **367**, 282–294 (1974). p. 141.
- [330] Reits, E. & Neefjes, J. From fixed to frap: measuring protein mobility and activity in living cells. *Nat Cell Biol* **3**, 145–147 (2001).
- [331] Sprague, B., Pego, R., Stavreva, D. & McNally, J. Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophys J* **86**, 3473–3495 (2004). pp. 143 and 165.
- [332] Mueller, F., Mazza, D., Stasevich, T. & McNally, J. Frap and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know? *Curr Opin Cell Biol* **22**, 403–411 (2010). p. 141.
- [333] Jacobson, K., Wu, E. & Poste, G. Measurement of the translational mobility of concanavalin in a glycerol-saline solutions and on the cell surface by fluorescence recovery after photobleaching. *Biochim Biophys Acta* **433**, 215–222 (1976). pp. 141 and 164.

- [334] Nuñez, D. *Thesis (S.B.) - Experimental estimate of the diffusivity of Vascular Endothelial Growth Factor*. Master's thesis, Dept. of Mechanical Engineering, Massachusetts Institute of Technology (2006). p. 141.
- [335] Vempati, P., Mac Gabhann, F. & Popel, A. Quantifying the proteolytic release of extracellular matrix-sequestered vegf with a computational model. *PLoS One* **5**, e11860 (2010). p. 141.
- [336] Czirok, A. Endothelial cell motility, coordination and pattern formation during vasculogenesis. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **5**, 587–602 (2013). pp. 145 and 151.
- [337] Murray, J. & Oster, G. Cell traction models for generating pattern and form in morphogenesis. *J Math Biol* **19**, 265–279 (1984). p. 145.
- [338] Vernon, R., Angello, J., Iruela-Arispe, M. L., Lane, T. & Sage, E. H. Reorganization of basement membrane matrices by cellular traction promotes the formation of cellular networks in vitro. *Lab Invest* **66**, 536–547 (1992). p. 145.
- [339] Donovan, D., Brown, N., Bishop, E. & Lewis, C. Comparison of three in vitro human "angiogenesis" assays with capillaries formed in vivo. *Angiogenesis* **4**, 113–121 (2001). p. 145.
- [340] Kleinstreuer, N. *et al.* A computational model predicting disruption of blood vessel development. *PLoS Comput Biol* **9**, e1002996 (2013). pp. 145 and 151.
- [341] Scianna, M. A multiscale hybrid model for pro-angiogenic calcium signals in a vascular endothelial cell. *Bull Math Biol* **74**, 1253–1291 (2012). p. 146.
- [342] Czirok, A., Zamir, E. A., Szabo, A. & Little, C. D. Multicellular sprouting during vasculogenesis. *Curr Top Dev Biol* **81**, 269–289 (2008). p. 147.
- [343] Bentley, K., Gerhardt, H. & Bates, P. A. Agent-based simulation of notch-mediated tip cell selection in angiogenic sprout initialisation. *J Theor Biol* **250**, 25–36 (2008).
- [344] Bentley, K., Mariggi, G., Gerhardt, H. & Bates, P. A. Tipping the balance: robustness of tip cell selection, migration and fusion in angiogenesis. *PLoS Comput Biol* **5**, e1000549 (2009). p. 147.
- [345] Zieris, A. *et al.* Fgf-2 and vegf functionalization of starpeg–heparin hydrogels to modulate biomolecular and physical cues of angiogenesis. *Biomaterials* **31**, 7985–7994 (2010). p. 147.
- [346] Anderson, S. M., Siegman, S. N. & Segura, T. The effect of vascular endothelial growth factor (vegf) presentation within fibrin matrices on endothelial cell branching. *Biomaterials* **32**, 7432–7443 (2011).
- [347] Anderson, S. M. *et al.* Vegf internalization is not required for vegfr-2 phosphorylation in bioengineered surfaces with covalently linked vegf. *Integrative Biology* **3**, 887–896 (2011).
- [348] Tsurkan, M. V. *et al.* Growth factor delivery from hydrogel particle aggregates to promote tubular regeneration after acute kidney injury. *J Controlled Release* **167**, 248–255 (2013). p. 147.
- [349] Box, G. E. Robustness in the strategy of scientific model building. Tech. Rep., DTIC Document (1979). pp. 149 and 152.
- [350] Khadra, A. & Schnell, S. Development, growth and maintenance of  $\beta$ -cell mass: Models are also part of the story. *Mol Aspects Med* **42**, 78–90 (2015). p. 150.

- [351] Gage, B. K., Webber, T. D. & Kieffer, T. J. Initial cell seeding density influences pancreatic endocrine development during in vitro differentiation of human embryonic stem cells. *PLoS one* **8**, e82076 (2013). p. 150.
- [352] Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat Methods* (2014). p. 150.
- [353] Clause, K. C. & Barker, T. H. Extracellular matrix signaling in morphogenesis and repair. *Curr Opin Biotechnol* **24**, 830–833 (2013). p. 151.
- [354] Daub, J. T. & Merks, R. M. A cell-based model of extracellular-matrix-guided endothelial cell migration during angiogenesis. *Bull Math Biol* **75**, 1377–1399 (2013). p. 151.
- [355] Méhes, E. & Vicsek, T. Collective motion of cells: from experiments to models. *Integrative Biology* **6**, 831–854 (2014).
- [356] Hayenga, H. N. *et al.* Multiscale computational modeling in vascular biology: From molecular mechanisms to tissue-level structure and function (2013).
- [357] Heck, T., Vaeyens, M.-M. & Van Oosterwyck, H. Computational models of sprouting angiogenesis and cell migration: towards multiscale mechanochemical models of angiogenesis. *Math Model Nat Phenom* (2015).
- [358] Korvasová, K., Gaffney, E., Maini, P., Ferreira, M. & Klika, V. Investigating the turing conditions for diffusion-driven instability in the presence of a binding immobile substrate. *J Theor Biol* **367**, 286–295 (2015). p. 151.
- [359] Moon, J. J. & West, J. L. Vascularization of engineered tissues: approaches to promote angiogenesis in biomaterials. *Curr Top Med Chem* **8**, 300–310 (2008). p. 151.
- [360] Auger, F. A., Gibot, L. & Lacroix, D. The pivotal role of vascularization in tissue engineering. *Annu Rev Biomed Eng* **15**, 177–200 (2013). p. 151.
- [361] Chwalek, K., Tsurkan, M. V., Freudenberg, U. & Werner, C. Glycosaminoglycan-based hydrogels to modulate heterocellular communication in in vitro angiogenesis models. *Science Reports* **4**, 4414 (2014). p. 151.
- [362] Montenegro, C. F. *et al.* Blocking  $\alpha v \beta 3$  integrin by a recombinant rgd disintegrin impairs vegf signaling in endothelial cells. *Biochimie* **94**, 1812–1820 (2012). p. 151.
- [363] Sluka, J. P. *et al.* The cell behavior ontology: describing the intrinsic biological behaviors of real and model cells seen as active agents. *Bioinformatics* **30**, 2367–2374 (2014). p. 153.
- [364] Maška, M. *et al.* A benchmark for comparison of cell tracking algorithms. *Bioinformatics* **30**, 1609–1617 (2014). p. 154.
- [365] Preibisch, S., Saalfeld, S., Schindelin, J. & Tomancak, P. Software for bead-based registration of selective plane illumination microscopy data. *Nat Methods* **7**, 418–419 (2010). p. 154.
- [366] Preibisch, S. *et al.* Efficient bayesian-based multiview deconvolution. *Nat Methods* (2014). p. 154.
- [367] Salbreux, G., Barthel, L. K., Raymond, P. A. & Lubensky, D. K. Coupling mechanical deformations and planar cell polarity to create regular patterns in the zebrafish retina. *PLoS Comput Biol* **8**, e1002618 (2012). p. 155.
- [368] Uriu, K., Ares, S., Oates, A. C. & Morelli, L. G. Dynamics of mobile coupled phase oscillators. *Physical Review E* **87**, 032911 (2013). p. 155.



- [369] Podgorski, G. J., Bansal, M. & Flann, N. S. Regular mosaic pattern development: a study of the interplay between lateral inhibition, apoptosis and differential adhesion. *Theor Biol Med Model* **4**, 43 (2007). p. 155.
- [370] Bentley, K. *et al.* The role of differential ve-cadherin dynamics in cell rearrangement during angiogenesis. *Nat Cell Biol* (2014). p. 155.
- [371] Salazar-Ciudad, I., Jernvall, J. *et al.* A computational model of teeth and the developmental origins of morphological variation. *Nature* **464**, 583–586 (2010). p. 155.
- [372] Jönsson, H., Heisler, M. G., Shapiro, B. E., Meyerowitz, E. M. & Mjolsness, E. An auxin-driven polarized transport model for phyllotaxis. *Proc Natl Acad Sci U S A* **103**, 1633–1638 (2006). p. 155.
- [373] Buske, P. *et al.* A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt. *PLoS Comput Biol* **7**, e1001045 (2011). p. 155.
- [374] Buske, P. *et al.* On the biomechanics of stem cell niche formation in the gut—modelling growing organoids. *FEBS J* **279**, 3475–3487 (2012). p. 155.



## VERSICHERUNG

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Diese Dissertation wurde an der Technischen Universität Dresden, Zentrum für Informationsdienste und Hochleistungsrechnen, Abteilung Innovative Methoden des Computing, unter der wissenschaftlichen Betreuung von Prof. Dr. Andreas Deutsch angefertigt.

Hiermit versichere ich, dass ich keine früheren erfolglosen Promotionsverfahren bestritten habe.

Hiermit erkenne ich die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften an der Technischen Universität Dresden in der Fassung vom 23. Februar 2011 an.

Datum, Unterschrift: