Collective features in the control of cell behavior

Dóra Lakatos

Supervisor:
ANDRÁS CZIRÓK
PhD, associate professor
Eötvös Loránd University
Department of Biological Physics

Eötvös Loránd University
Faculty of Science
Doctoral School of Physics
Statistical physics, biological physics and physics of quantum systems program

Head of doctoral school: Prof. Dr. Tamás Tél
Head of doctoral program: Prof. Dr. Jenő Kürti

2018
# Contents

## 1 Objectives

## 2 Introduction

2.1 Receptors .......................... 4

2.2 Extracellular signaling .......................... 6

2.2.1 Juxtacrine signaling ....................... 6

2.2.2 Autocrine signaling ....................... 6

2.2.3 Paracrine signaling ....................... 7

2.2.4 Endocrine signaling ....................... 8

2.3 Regulatory networks ....................... 8

## 3 Autocrine feedback in pluripotent stem cell maintenance

3.1 Background .......................... 11

3.2 Methods .......................... 13

3.2.1 Mouse ES Cell Culture ....................... 13

3.2.2 Fluorescent Cell Sorting and Analysis ....................... 13

3.2.3 RNA Analysis .......................... 13

3.2.4 Protein Analysis .......................... 14

3.2.5 Dox-inducible Nanog ES cell line ....................... 14

3.3 Results .......................... 14

3.3.1 Model structure ....................... 14

3.3.2 The core network ....................... 21

3.3.3 Sensitivity of the core network to model parameters ....................... 24

3.3.4 NANOG heterogeneity ....................... 26

3.3.5 KLF4 and ESRRB ....................... 29

3.3.6 Autocrine FGF signaling ....................... 29

3.3.7 NANOG expression determined by autocrine feedback ....................... 31

3.4 Discussion .......................... 34

3.4.1 The mechanism behind Nanog fluctuations ....................... 34

3.4.2 Comparison with experimental data ....................... 36

3.4.3 Future directions ....................... 39

3.5 Conclusions .......................... 39

## List of Figures

v

## List of Tables

vii
List of Figures

3.1 Suggested interactions of the NANOG-OCT4-SOX2 core module . . . . . . 12
3.2 Studied NANOG core circuit models . . . . . . . . . . . . . . . . . . . . . . 22
3.3 Steady state properties of the NANOG core circuit model variants . . . . . 25
3.4 NANOG circuit parameter sensitivity analysis . . . . . . . . . . . . . . . . . 26
3.5 Transcriptional activity during NANOG fluctuations . . . . . . . . . . . . . 27
3.6 Proposed ES cell maintenance signaling network . . . . . . . . . . . . . . . 28
3.7 KLF4-ESRRB module as an amplifier . . . . . . . . . . . . . . . . . . . . . . 29
3.8 Modulating the transcriptional regulation of \textit{Nanog} . . . . . . . . . . . . . 32
3.9 Substates within the “ON” state of the core \textit{Nanog} switch . . . . . . . 32
3.9 Substates within the “ON” state of the core \textit{Nanog} switch (cont.) . . . . 33
3.10 Noise driven fluctuations in protein concentrations and gene expression . 35
4.1 Model of sVEGFR1 driven vascular pattern formation . . . . . . . . . . . . . 47
4.2 Sprouting assay using HUVEC spheroids in fibrin gel . . . . . . . . . . . . . 54
4.3 HUVEC spheroid after 6 days in culture . . . . . . . . . . . . . . . . . . . . . 55
4.4 Effect of sVEGFR1 on sprout density . . . . . . . . . . . . . . . . . . . . . . 56
4.5 Cumulative distribution of radial density profiles . . . . . . . . . . . . . . . 57
4.6 Time development of the lattice model . . . . . . . . . . . . . . . . . . . . . 59
4.7 Time development of the phase-field model . . . . . . . . . . . . . . . . . . . 60
4.8 Time development of the lattice model using rescaled parameter set . . . . . 61
4.9 Tip-splitting event in the phase-field model . . . . . . . . . . . . . . . . . . . 62
4.10 VEGF concentration profiles along a growing branch tip . . . . . . . . . . . 62
4.11 Time development of the diffusing VEGF-driven phase-field model . . . . . 63
4.12 Dense branching morphology of the lattice model . . . . . . . . . . . . . . . 64
4.13 Parameter sensitivity of branching morphology in the lattice model . . . . . 64
4.14 Parameter sensitivity of branching morphology in the phase-field model . . 65
4.15 Representative sprout density profiles in the computational models . . . . 65
4.16 Velocity selection in the lattice and phase-field models . . . . . . . . . . . . 69
4.17 The lattice model with heterogeneous sVEGFR1 secretion . . . . . . . . . . 72
4.18 Patterning with heterogeneous sVEGFR1 production . . . . . . . . . . . . . 73
4.19 Patterning in the presence of VEGF internalization . . . . . . . . . . . . . . 74
5.1 Space-time plot obtained with the logistic growth term . . . . . . . . . . . . . 85
5.2 Allee effect in glioblastoma cell cultures . . . . . . . . . . . . . . . . . . . . . 86
5.3 Space-time plot obtained with Allee effect . . . . . . . . . . . . . . . . . . . . 88
5.4 Total number of cancer cells for different values of the resection threshold . 88
5.5 Remission time for weak and strong Allee effects . . . . . . . . . . . . . . . . 89
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Nanog model parameters I</td>
<td>20</td>
</tr>
<tr>
<td>3.2</td>
<td>Nanog model parameters II</td>
<td>20</td>
</tr>
<tr>
<td>3.3</td>
<td>Nanog model parameters III</td>
<td>20</td>
</tr>
<tr>
<td>4.1</td>
<td>VEGF model parameters</td>
<td>66</td>
</tr>
</tbody>
</table>
Chapter 1

Objectives

A major challenge of cell biology is to understand how cellular processes, such as cell survival, proliferation, growth or differentiation, are regulated by a complex network of molecular agents. The systems biology approach tries to help our understanding by constructing mathematical or computational models and use experimental results to validate these models. Elucidating the complex regulatory dynamics underlying collective cell behavior and tissue function requires examination of multiple factors and components acting on various spatial and time ranges, that is presently unfeasible. However, analyzing the contribution of a subset of regulatory mechanisms to an emergent (cell- or tissue-level) phenomena can provide valuable insights into the working principles of molecular control mechanisms. In this work, we focus on three related problems, where collective behavior emerges through feedback regulatory mechanisms acting on various time and length scales: the decision mechanism involved in stem cell maintenance (Chapter 3), the role of diffusive signal inhibitors in vascular patterning (Chapter 4), and the role of cellular cooperative behavior in tumor recurrence after resection (Chapter 5).

The maintenance of stem cell pluripotency is controlled by a core cluster of transcription factors, NANOG, OCT4 and SOX2 – genes that jointly regulate each other’s expression. The expression of some of these genes, especially of Nanog, is heterogeneous in a population of undifferentiated stem cells in culture. Transient changes in expression levels, as well as heterogeneity of the population, is not restricted to this core regulator, but involve a large number of other genes that include growth factors, transcription factors or signal transduction proteins. Since the molecular mechanisms behind NANOG expression heterogeneity are not yet understood, we explore by computational modeling the core transcriptional regulatory circuit and its input from autocrine FGF signals that act through the MAP kinase cascade. We demonstrate that autocrine feedback regulatory loops are expected to result in transcriptional states (attractors) with distinct expression profiles. We also measure expression levels of
transcription factors in heterogeneous stem cell populations, and compare them to the computational model predicted results.

Vascular patterning is a key process during development and disease, that emerges from the collective behavior of endothelial cells, that likely utilizes a variety of guidance mechanisms. One of the best established regulators of vascular growth is VEGF, that can be sequestered in the ECM and induce cell motility, proliferation and chemotaxis. However, pre-patterned VEGF in the ECM has not been demonstrated in many vascular patterning phenomena. Recent experiments demonstrated that secreted type 1 VEGF receptors (sVEGFR1) can modulate endothelial cell behavior, yet the mechanism by which it controls vascular structure is little understood. We propose computational models to shed light on how vascular patterning is guided by self-organized gradients of the VEGF/sVEGFR1 factors, and elucidate whether a diffusive inhibitor can generate structures with a dense branching morphology in models where the activator elicits directed growth. We compare model predictions with time-resolved experimental data obtained from endothelial sprout kinetics in fibrin gels.

In the case of solid cancers, such as glioblastoma, primary therapy often involves the surgical removal of the tumor. Resection of the bulk of a tumor often cannot eliminate all cancer cells, due to their infiltration into the surrounding healthy tissue, that may lead to recurrence of the tumor at a later time. At the macroscopic level, tumor progression can be described by mathematical models specifying the spatiotemporal changes in tumor cell density. In combination with diagnostic imaging, such predictive quantitative models can provide patient-specific computational optimization in treatment strategies. The interaction between tumor cells, mediated by soluble diffusing factors, may profoundly affect their migratory and proliferative activity. To gain a qualitative insight into the density dependent proliferation and motility of cancer cells, we perform experiments in vitro with glioblastoma cell lines, particularly at low cell densities. To understand how a cell density-dependent proliferation affects the recurrence of tumor growth, we investigate front propagation in a generalized Fisher-Kolmogorov equation.
Chapter 2

Introduction

Cellular signaling is a crucial prerequisite for the existence of multicellular organisms, as cells in such an organism must regulate their behavior (proliferation, motility, etc.) in response to internal and external stimuli. The active communication between cells and their environment enables the emergence of collective behavior: much more complex responses are possible (such as development of functional tissues, or healing in response to injury) than the individual constituents are capable of [1–5]. Cellular information transduction is mediated through nonlinear biochemical networks, that are typically interconnected and multifunctional [2]. Furthermore, they are exposed to intrinsic and extrinsic fluctuations, giving rise to heterogeneity even in isogenic cell populations [5–8]. Hence, signaling networks must operate in a robust and adaptive manner to ensure the maintenance of cellular identity and proper tissue function. Elucidating the regulatory mechanism underlying collective cell behavior and tissue function is still a challenging problem in systems biology [3, 6].

Cellular response to external stimuli is initiated from cell membrane receptors, that can be activated by various extracellular ligands [9]. Extracellular signaling molecules can act over various spatial and time ranges, encompassing several orders of magnitude by utilizing a multitude of regulating mechanisms [10]. While some receptors, such as G protein-coupled receptors (GPCRs) and receptor tyrosine kinases (RTKs), bind soluble factors, others can be activated by binding extracellular matrix (ECM) proteins or adjacent cell surface receptors (through adhesion and cell-cell contact) [2]. Upon activation, changes occur in receptor conformation, that generate a cascade of events within the cells [11]. The intracellular signaling pathways involved in such changes utilize a handful of molecular mechanisms and often relay the external signals to the regulatory apparatus of gene transcription and translation [12, 13]. Alterations in gene expression profiles change protein concentrations, therefore capable of modifying the propagation of future signals. Thus, feedback loops and cross-talks
arise within signaling and gene regulatory networks, whose functions are often challenging to understand [2, 14]. Still, such interactions enable the emergence of specific cellular behaviors, such as proliferation, differentiation or apoptosis [11, 13, 15].

2.1 Receptors

Cell surface receptors mediate responses to diverse extracellular signaling molecules. In addition to the transmembrane regions, most receptors contain two functional domains: an extracellular ligand binding domain and an intracellular effector domain. The separation of this two functions allows the receptor to produce different intracellular signals for diverse ligands. Receptor domains can exist in multiple conformations, that can be either active or inactive in the aspect of signaling. Interaction with the ligand can promote or inhibit the shift in conformational equilibrium toward active conformations, or can initiate oligomerization of receptors resulting in autophosphorylation. Ligand binding affinity and signal generating capability of the receptor can be altered by allosteric regulation, also known as cooperativity, and by covalent modifications, e.g. phosphorylation or ubiquitination. Sensitivity to an extracellular signal can also be controlled by cell surface receptor density, that is also a subject to diverse regulating mechanisms acting on different time scales. Receptors typically initiate the activation of multistep signaling pathways and act as molecular amplifiers through their intrinsic enzymatic activities or indirectly, through recruitment of intracellular enzymes and effector proteins. Transactivation between different type of receptors is also a known phenomenon, that plays an important role in signal transduction [16].

Most cell surface receptors can be divided into three classes based on the transduction mechanism utilized. Ion channel receptors are usually oligomers formed by homologous units containing several transmembrane regions, that allow passive diffusion of ions across the cell membrane in a selective manner. Ion channels can exist in open and closed conformations, transient opening can occur upon ligand binding or perception of electrical impulses, while channel closing is typically an automatic mechanism. In multicellular organisms, functions of ion channels include the maintenance of resting membrane potential, and generation of action potentials and synaptic responses in electrically excitable cells [11].

G protein-coupled receptors comprise a large receptor family with diverse functions, all composed of seven transmembrane $\alpha$-helices. GPCRs can be activated by binding of various types of extracellular ligands, that induce conformational changes in the receptor. Since GPCRs lack the intrinsic enzymatic domain, they act indirectly
on their effector proteins by promoting the activation of heterotrimeric GTP-binding proteins (G proteins). Membrane-associated G proteins then can convey the signal to several effector proteins, that initiate intracellular signaling cascades through their enzymatic activity. Due to its modularity and variability in the level of signaling components, GPCR signaling is involved in the regulation of most cell functions [11, 16].

Enzyme-linked receptors are single-pass transmembrane helices with an extracellular ligand binding domain and a cytoplasmic domain, that either has an intrinsic catalytic activity or associates directly with an enzyme, and generally form dimers or oligomers. This family comprises numerous receptors with heterogeneous structures and distinct intracellular enzymatic functions, such as protein kinase, phosphatase or cyclase activity. The great majority of them are receptor tyrosine kinases, that can phosphorylate the tyrosine residue of their target protein upon activation. Most RTKs contain phosphorylation sites in their cytoplasmic domain, that are autoinhibited in their inactive, typically monomeric state. In the canonical view, ligand binding induces dimerization or oligomerization of RTKs either in a homo- or heterotypic way, that catalyzes auto- or trans-phosphorylation of the receptors’ distinct residues in the cytoplasmic domain. By promoting conformational changes, autophosphorylation enables the creation of docking sites for the recruitment of effector proteins, that initiate the activation of intracellular signaling cascades [11, 17–19]. However, an increasing number of studies suggest, that many RTKs can be primed by forming dimers or oligomers even in the absence of the ligand. In both cases, modulation of lateral receptor distribution in the cell membrane is a key rate-limiting step of RTK activation [18, 19].

Extracellular ligands of RTKs are primarily growth factors, that can bind their target receptor with high affinity and specificity and may also form homo- or heterodimers. Many growth factors have multiple isoforms, differing in their spatiotemporal dynamics and bioactivity due to their different physical and biochemical properties, especially their diffusivity and ECM-binding capability. Furthermore, these isoforms may also trigger distinct cellular response and collective behavior, such as vascular endothelial growth factor (VEGF) isoforms in angiogenesis [20–22]. Signaling through RTKs is particularly mediated in an autocrine and a paracrine manner and regulates diverse cellular functions, deregulation of them is critically involved in the development and progression of many cancer types and diseases [19, 23].
2.2 Extracellular signaling

2.2.1 Juxtacrine signaling

Juxtacrine signaling is mediated through different cellular adhesion molecules, that can selectively bind proteins and oligosaccharides tethered to neighboring cells or to the extracellular matrix, therefore enabling cells to transduce chemical and mechanical signals. Due to the nondiffusible nature of both ligand and receptor, juxtacrine signaling is spatially restricted to the close proximity of the cell, being essential in multicellular processes from cell fate decision to pattern formation [11, 24, 25].

Cell adhesion can operate on different time scales. Transient cell adhesions are typically mediated by integrins, Notch signaling, selectins and immunoglobulins, and play a key role in developmental processes, cell migration and immune and inflammatory responses. Transient cell-cell or cell-ECM contacts may be stabilized by junction formation, that involves the reorganization of the cytoskeleton and tight linkage of it to adhesion receptors by cytoplasmic anchor proteins. Such stable cell-ECM adhesions are indispensable for cell migration and mainly regulated by integrins [11, 24]. Formation of stable cell-cell junctions is mediated by cadherins, that are key regulators of cell recognition and the establishment and maintenance of tissue homeostasis, with diverse effects on cell behavior [11, 26].

Adjacent cells can also form direct connections between their cytoplasms through gap junctions, that allow the exchange of small intracellular mediators, such as ions, short peptides and second messengers, by diffusion. This direct and rapid intercellular communication is important in many physiological processes, e.g. smooth and cardiac muscle contraction, neuronal excitability or epithelial electrolyte transport [11, 27]. Cells tune their juxtacrine signaling efficiency through several regulatory mechanisms, such as ligand and receptor post-translational modifications mediating exocytosis, membrane trafficking, binding affinity and endocytosis [11, 28].

2.2.2 Autocrine signaling

Extracellular signaling can also be mediated through secreted molecules, that can interact with cell membrane receptors and extracellular matrix components or be carried through the circulation that can occur at various distances [24]. In case of autocrine signaling, cells synthesize and secrete soluble ligands that activate receptors on their own surfaces, thus creating a feedback loop through intracellular signal transduction [29]. These signaling circuits can operate both as positive and negative regulators,
resulting in intrinsically self-contained loops [30]. While their experimental investigation is difficult due to their recursive and spatially restricted nature, quantitative approaches could shed light on their operational principles [29, 31]. In particular, receptor binding and diffusive loss of secreted ligands are the two competing processes that are mediated by multiple kinetic and biochemical parameters, such as the ligand secretion rate, ligand diffusivity, the receptor-ligand binding rate, the internalization and degradation rates of the ligand and receptors, or receptor surface density. [29, 31, 32]. Within cell populations, autocrine signaling activity can also depend on cell density, cell developmental state and tissue architecture, hence it is a regulatory mechanism that can give rise to complex spatiotemporal patterns [11, 33, 34]. Accordingly, autocrine loops are important regulators of stem cell pluripotency, embryonic development, tissue homeostasis and they are also involved in many pathological conditions, such as cancer and autoimmune diseases [31, 35].

Signaling mediated through the epidermal growth factor receptor (EGFR) is one of the best understood autocrine signal transduction pathways. It is activated by its autocrine ligands, such as epidermal growth factor (EGF) or transforming growth factor alpha (TGF-α), and regulates cellular processes that include cell division, differentiation, migration and polarity [34, 36–38]. Upon EGFR activation, the cellular response is particularly mediated through the Ras-MAPK (small GTPase, mitogen-activated protein kinase) pathway, and can modulate the presence of ligands and receptors on the cell surface. This modulation of ligand secretion and receptor availability constitutes a positive feedback [31, 39]. Abnormal amplification of such EGFR activity has been identified in multiple cancer types, e.g. glioblastoma, metastatic colorectal cancer and non-small-cell lung cancer [40]. Furthermore, Ras-driven upregulation of EGFR autocrine signaling has been found to enhance radiation resistance in cancer cells [41].

2.2.3 Paracrine signaling

In paracrine signaling, cell secreted ligands can interact with ECM components and other extracellular molecules before binding to surface receptors of some other cell type. The diffusible ligands can thus typically exert their effect in a range of several cell diameters [10, 42, 43]. Similar to autocrine signaling, spreading of the secreted ligand is controlled by several interdependent factors resulting in diverse spatiotemporal patterns [2, 10]. The concentration profile of the signaling molecule is primarily regulated by ligand secretion and degradation dynamics, and physicochemical properties of the ligand and the extracellular space. While diffusion of soluble proteins is commonly hindered owing to hydrodynamic interactions with the ECM, ligand’s
spreading distribution may also be modulated by reversible binding to ECM components, e.g. heparan sulfate proteoglycans (HSPGs) [2, 44]. Such retention or immobilization of the ligand can significantly alter its effective diffusivity and lifetime as well, by preventing it from physical denaturation, degradation by extracellular proteases and cellular uptake [2, 10, 44, 45]. Long-range spreading of insoluble or large proteins is often facilitated by active vesicular transport mechanism, that shuttles the ligand through the cytoplasm of other cells in a spatially-directed manner [2, 46].

Cells are able to sense both concentration levels and gradients of secreted ligands, and their response – that could be both dose-dependent or switch-like – can involve changes in gene expression patterns [12, 47, 48]. Thus, precise and robust regulation of ligand distribution is indispensable in the fluctuating environment for proper signaling, which is achieved through diverse extracellular feedback mechanisms, such as enhanced degradation or graded activation profile [49, 50]. Formation of concentration gradients by secreted molecules drives numerous morphological processes during development and in adult tissues, such as vasculogenesis and angiogenesis [11, 51, 52]. Most patterning processes emerge from a dynamic cross-talk between paracrine and autocrine morphogen gradients and mechanical forces that cooperatively guide cellular motility and cell fate decisions [52, 53].

2.2.4 Endocrine signaling

Hormones, the messenger molecules of endocrine signaling, are secreted by specialized endocrine cells and carried by the circulatory system to act on distant targets [11]. Hence, hormones must be synthesized in larger amounts and operate on slower timescales than autocrine and paracrine signaling molecules. Furthermore, hormones might require specific post-translational modifications against extracellular degrading enzymes and reversible binding by carrier proteins for effective transportation in large concentrations [54]. Depending on the hormone’s chemical structure, it can activate cell surface receptors, such as GPCRs or RTKs, or traverse through lipid membranes and bind intracellular receptors, thus initiating downstream signaling [55].

2.3 Regulatory networks

Cellular behavior is tightly controlled by a number of interacting molecular networks that involve cellular messenger substances like ions and small organic molecules, post-translational modifications of signaling proteins (signal transduction) and regulation of gene transcription and translation [4, 56]. These regulatory networks often
thought to consist of functional modules and operate in a nonlinear and hierarchical manner, often localized to a certain cellular compartment and giving rise to both robustness and flexibility at the same time [2, 57]. Despite the diversity of sensed substances, the number of biochemical components and mechanisms involved in signaling pathways are typically conserved and reused [11, 16, 58].

Since receptors often have overlapping activation patterns, further integration of incoming stimuli in accordance with internal signals and transduction of it to the nucleus is required for proper cellular function. This complex mechanism is primarily achieved by the use of multistep and highly interconnected networks consisting of small signaling molecules and intermediary proteins, that give rise to signal processing algorithms, such as signal amplification and routing to different targets, spatial and temporal regulation of signal kinetics, feedback loops, or insertion of control points. Many signaling proteins contain multiple modules, that may possess enzymatic functions or enable specific protein-protein interactions. Complex formation in particular, either being stable or transient, is mediated through a variety of binding domains. Assembly of them can enhance signaling efficiency by increasing local concentration of interacting components, promoting low affinity binding or facilitating activation or inhibition [11, 16].

Different spatiotemporal activation profiles of the same kind of signaling molecules can also result in distinct cellular responses. The most understood example of this behavior is the activation dynamics of MAPK/ERK (extracellular signal-regulated kinase) pathway through RTKs in neural progenitor cells (PC12). While treatment of these cells with fibroblast growth factor (FGF) or nerve growth factor (NGF) leads to sustained activation of ERK and induces differentiation of the cells, stimulation with EGF generates transient ERK phosphorylation profiles and results in proliferation. The distinct MAPK activation dynamics and cellular responses emerge from an intricate intracellular regulation of signal transduction, a mechanism that also involves feedback loops through gene regulatory networks. Since MAPKs can directly regulate their transcription factor substrates in the nucleus, their spatial distribution in different compartments also plays a critical role in cellular response generation [12, 59–61].

A very important possible component of the intracellular biochemical signal processing mechanism is the alteration of gene expression profiles, thus alteration in the biochemical makeup of the cell and possibly the operational logic of its signal processing apparatus. In particular, distinct cellular phenotypes and functions can be thought of as attractor states of gene expression patterns, that arise from dynamic interactions between genes and their regulatory proteins [13]. Gene expression is primarily limited
by its transcription, that is regulated by transcription factors (TFs), RNA polymerase (RNAP) and components of the transcriptional machinery. General TFs can bind the promoter region of the gene and form the preinitiation complex, that can recruit RNA polymerase and initiate gene transcription. DNA sequence-specific TFs can also bind the regulatory gene site and act as an enhancer or a repressor, by increasing or decreasing its transcription rate through cooperative interactions with the initiation complex and other cofactors [11, 62]. Regulatory mechanisms of transcription factor activation can be diverse and occur on different time ranges. While TFs already located in the nucleus can be activated or deactivated in response to external and internal stimuli through second messengers or activation of nuclear receptors or by translocated signaling proteins, cytoplasm located TFs can be activated and delivered into the nucleus in response to external signals [63].
Chapter 3

Autocrine feedback in pluripotent stem cell maintenance

3.1 Background

Embryonic stem (ES) cells are pluripotent cell populations that can be induced to differentiate into a variety of cell types. Mouse ES cells are derived from the inner cell mass of the blastocyst, and their capacity to either self-renew or differentiate into cells of the three germ layers; thus this cell type is a useful model to dissect the molecular regulation behind pluripotency and differentiation [64–66]. The maintenance of stem cell pluripotency is controlled by a core cluster of transcription factors (TF), including NANOG, OCT4 and SOX2. The molecular mechanisms by which these factors act is complex and not completely characterized; however part of their critical activity includes the joint regulation of each other’s expression [67, 68]. The pluripotent status of stem cells is maintained through high expression levels of these genes, and the downregulation of these factors accompanies and is required for cell differentiation. NANOG expression levels are crucial as its forced expression is sufficient to sustain pluripotency even in the absence of extracellular signaling factors such as Leukemia Inhibiting Factor (LIF) [69, 70] which are otherwise required for stem cell maintenance.

Since the groundbreaking analysis of Chickarmane et al. [71], the structure of the NANOG-OCT4-SOX2 transcriptional regulatory network has been revised – the recently proposed models are compared in Fig. 3.1. According to the current consensus, the OCT4/SOX2 dimer acts as a common transcription factor for all three genes and no autocatalytic activation of NANOG takes place. High OCT4 levels were suggested to be repressors of NANOG (in addition to the activator function of the OCT4/SOX2 dimer) [68, 73]. More recently, the OCT4 inducing function of NANOG was questioned, while NANOG was suggested to act as an autorepressor [72].
In the last few years it also became clear that – despite their fundamental importance – the expression of many components of the stem cell self-renewal circuitry, including NANOG, is heterogeneous in a population of undifferentiated stem cells [70, 73, 74]. The heterogeneity is dynamically maintained, with individual cells exhibiting transient changes in expression levels. Genes with dynamic expression in mouse ES cells include growth factors, transcription factors, and signal transduction proteins [72, 75, 76]. Recently Galvin-Burgess et al. proposed that undifferentiated ES cells are in various distinct states, depending on the activity of various signal transduction pathways [76]. The presence of states with distinct Nanog expression levels was also suggested based on statistical modeling of changes in flow-sorted populations [77]. Conceptually, the pluripotent and differentiating states of these cells are thus not described well by a simple “ON/OFF” switch. Instead, a cell being in one of the various pluripotent states may be primed or biased in a way that influences its response to differentiation-inducing signals [78].

In view of these developments, we revisit the dynamics of the core NANOG transcriptional regulatory circuit. As shown in Fig. 3.2, we will consider the OCT4/SOX2 dimer as a common transcription factor for all three genes, and the NANOG protein to be a transcription enhancer for the Sox2 gene. We consider four model scenarios, in which NANOG either is or is not an inducer of Oct4. We consider the models proposed by Pan et al. where high OCT4 levels are repressors of Nanog and Oct4 [68], and that of Navarro et al., which includes an autorepressor feedback to Nanog [72]. By numerical simulations we demonstrate that all these models result in a bistable, switch-like behavior. To address the observed heterogeneity in Nanog expression levels, we also explore a biologically plausible scenario to couple the core circuit to extracellular signals. Based on simulation results we argue that instead of an instability within the
core regulatory circuit, fluctuations in Nanog expression levels and associated distinct cell states are likely to be generated by stochastic autocrine feedback loops, like the one involving secreted FGFs.

3.2 Methods

3.2.1 Mouse ES Cell Culture

Experiments used the BAC-Nanog-GFP (BNG) ES cells [76] in which Nanog-GFP bacterial artificial chromosome (BAC) [79] was introduced into Ainv15 ES cells [80]. BNG ES cells were maintained as described previously [76, 81] on gelatin-coated or fibroblast cocultured plates. Cells were grown in serum-based ES cell media: Dulbecco’s modified Eagle’s medium, 15% fetal bovine serum (FBS), penicillin-streptomycin, L-glutamine, nonessential amino acids, β-mercaptoethanol, and $10^3$ units/ml leukemia inhibitory factor (LIF).

3.2.2 Fluorescent Cell Sorting and Analysis

ES cells were trypsinized to a single cell suspension and analyzed by BD FACSArray (cell sorting) and BD LSRII flow cytometer (cell analysis) for GFP expression. A convention was established for sorting and analyzing subpopulations of BNG ES cells based on the profile of GFP expression of unsorted cells, as described in [76]. The GFP-medium and GFP-high populations were determined by gating 30-35% of the cells from the peak of the GFP distribution. Cells expressing higher levels of GFP than the GFP-high subpopulation were classified as GFP-very high; whereas cells with expression less than GFP-medium cells but above background were identified as GFP-low cells. Analysis of control E14 ES cells were used as a negative control for flow analysis and sorting. After cell sorting, subpopulations were analyzed to confirm purity of the population.

3.2.3 RNA Analysis

RNA was isolated (Qiagen, Hilden, Germany), and cDNA was synthesized (Invitrogen) following manufacturer’s instructions. TaqMan primer sets with the 7500 Real Time polymerase chain reaction (PCR) system (Applied BioSystems, Foster City, CA) were used for quantitative real-time PCR analysis. Upon request, specific ABI TaqMan Primer/Probe assay identification numbers are available.
3.2.4 Protein Analysis

Cells were pelleted and then lysed with radioimmune precipitation assay buffer supplemented with Halt Protease and Phosphatase Inhibitor Cocktails (Pierce, Rockford, IL). Protein samples were separated on BioRad Tris-HCl gels, and blots were probed with primary antibodies for GAPDH (Santa Cruz Biotechnology, Santa Cruz, CA) and pERK (Cell Signaling) and incubated overnight at 4°C with appropriate secondary antibodies. SuperSignal West Pico Chemiluminescence (Pierce) was used to detect the Western blots.

3.2.5 Dox-inducible Nanog ES cell line

The Nanog coding sequence was inserted into a dox-responsive locus via CRE-mediated insertion using previously described methods [76, 80, 81] into the AInv-15 ES cell line, to generate BNG-dox-Nanog cell line. These cells were cultured under feeder-free conditions in standard serum based ES cell media supplemented with doxycycline (1 µg/ml) for the indicated times. Cells were then harvested in RIPA buffer for western analysis.

3.3 Results

3.3.1 Model structure

To explore the NANOG transcriptional regulatory network, we adopt the method of Chickarmane et al. [71]. The production and degradation of proteins are assumed to be much slower than the assembly or dissociation of multimolecular complexes, we thus include the latter processes using a quasi steady state approximation (QSSA).

With these simplifications the system is governed by three differential equations of the form

\[
(1 + D_G) \frac{d[G]}{dt} = \alpha_g p_g - \delta_G [G],
\]

where \( G \in \{\text{NANOG, OCT4, SOX2}\} \) are transcription factor proteins, \( g \) denotes the regulatory site of a gene \( G \), \( p_g \) is the probability of RNA Polymerase II (P) binding to the promoter \( g \); \( \alpha_g \) is the combined translation and transcription rate, and \( \delta_G \) is the decay rate of the proteins. The quasi steady state approximation yields the amount of complex-bound specimen \( [G_{\text{bound}}] \) as a function of the concentration of free specimen.
As a change in the total amount alters both the amount of free and complex-bound specimen,

\[ d[G] + d[G_{\text{bound}}] = d[G] \left( 1 + \frac{\partial[G_{\text{bound}}]}{\partial[G]} \right), \quad (3.2) \]

thus,

\[ D_G = \frac{\partial[G_{\text{bound}}]}{\partial[G]}, \quad (3.3) \]

As we show in the next subsection (Promoter activity), \( p_g \) can be written in the form of

\[ p_g = \frac{Z_{g}^{\text{ON}}}{Z_{g}}, \quad (3.4) \]

where

\[ Z_{g} = Z_{g}^{\text{ON}} + Z_{g}^{\text{OFF}}, \quad (3.5) \]

and the \( Z_{g}^{\text{ON}} \) and \( Z_{g}^{\text{OFF}} \) quantities are proportional to the probability of RNA polymerase II being bound or absent at locus \( g \), respectively. If the transcription logic is limited to two factors (enhancers or repressors) per locus, then for each locus \( g \in \{\text{Nanog, Oct4, Sox2}\} \) we obtain

\[ Z_{g}^{\text{ON}} = [P]K_{g,p} \left[ 1 + \sum_{H}[H]K_{g,H} \exp \left( \frac{C_{g,H}}{RT} \right) + \sum_{H,I}[H]K_{g,H}[I]K_{g,I} \exp \left( \frac{C_{g,H,I}}{RT} \right) \right], \quad (3.6) \]

and

\[ Z_{g}^{\text{OFF}} = 1 + \sum_{H}[H]K_{g,H} + \sum_{H,I}[H]K_{g,H}[I]K_{g,I}, \quad (3.7) \]

where \( H,I \in \{\text{NANOG, OCT4/SOX2, OCT4}\} \) are transcription factors bound to the locus \( g \). In the expressions (3.6) and (3.7) the equilibrium constant (binding affinity) of factor \( H \) to the binding site at locus \( g \) is denoted by

\[ K_{g,H} = \frac{\left[ G_{\text{bound}} \right]_{g}}{\left[ G_{\text{empty}} \right]_{[H]}}, \quad (3.8) \]

As we discuss in the next subsection, the equilibrium constants \( K \) as well as the cooperativity measures \( C \) are related to the binding energies between the transcription factors, the promoter and the RNA polymerase.
3.3.1.1 Promoter activity

We adopt the thermodynamic gene regulatory model proposed by Shea and Ackers [82], and later extended by Buchler et al. [83]. Their approach provides combinatorial modulation of the promoter by the polymerase and multiple transcription factors regardless of molecular mechanisms, by assuming that gene regulation is controlled by the equilibrium binding of the proteins to DNA and to each other [82–84].

Let us denote the Gibbs free energy of binding between a transcriptional regulator H and the corresponding regulatory site of gene g by \( B_{g,H} \). This quantity is related to the equilibrium constant \( K_{g,H} \) as

\[
RT \ln K_{g,H} c_0 = B_{g,H},
\]

(3.9)

where \( T \) is the temperature, \( R \) is the universal gas constant and \( c_0 \) is the standard concentration (the dimension of \( K \) being \( M^{-1} \)).

Following [82, 83], gene transcription activity is controlled by the binding probability of RNAP II (P), which, in turn is determined by the binding energy of the assembled complex at the promoter and regulatory sites. The binding energy of the complex, in the presence of a transcriptional regulator, H, is \( E_{g,H} \). It is convenient to express this quantity as

\[
E_{g,H} = B_{g,P} + B_{g,H} + C_{g,H},
\]

(3.10)

where the “cooperativity” \( C \) measures the difference between the binding energy of the complex and the binding energies of the individual constituents to the DNA sequence (in the absence of any other factors). In the simplest case \( C_{g,H} \) represents a “direct” binding energy between the transcription factor and RNAP II, but it can also reflect conformation changes induced by the presence of H. Similarly, in the case of a RNAP II complex containing two transcription factors, H and I, the binding energy can be written as

\[
E_{g,H,I} = B_{g,P} + B_{g,H} + B_{g,I} + C_{g,H,I}.
\]

(3.11)

The concentration or probability of a particular complex can be derived from the binding energies by the generalization of Eq. (3.9). For example, the concentration of the four component complex consisting the locus g, RNAP II and factors H and I is

\[
[\text{complex}] = [g][H][I][P]K_{g,H,I,P} = \frac{[g][H][I][P]}{c_0^3} \exp \left( \frac{E_{g,H,I}}{RT} \right),
\]

(3.12)

where \([g],[H],[I]\) and \([P]\) denotes the (time averaged) steady state concentration of
3.3. Results

the regulatory site \( g \) without any of the other model components bound to it, the concentrations of free transcription factors \( H \) and \( I \) and the concentration of unbound RNAP II, respectively.

### 3.3.1.2 Promoter binding parameters

To evaluate the promoter activity (Eq. (3.4)), the binding energies \( B \) and the \( C \) parameters characterizing cooperativity needs to be specified. As these values are not characterized in the literature, the following assumptions are made.

If transcription factor binding affinities are in the nanomolar range, Eq. (3.9) predicts a typical value of \( B = 12 \) kCal/mol. If intramolecular bonds between the components of the RNAP II complex is the main contributor to its enhanced stability, then a rough estimate for the \( C \) parameters is 2 kCal/mol for each pair of physical interaction [83], thus \( C = 4 \) kCal/mol for two transcription factors that each interact with the RNAP II but do not bind each other directly. Repression of transcription activity can be represented by large negative \( C \) values, which destabilizes the repressor-containing complex.

In the absence of any transcriptional regulator, the promoter activity is

\[
 p^{(0)}_g = \frac{[P]K_{g,P}}{1 + [P]K_{g,P}}. \tag{3.13}
\]

Thus the value of \([P]K_{g,P}\) can be estimated as the transcriptional activity without the considered transcription factors present. Knowing that in the absence of NANOG, OCT4 and SOX2 protein the transcription of these genes shuts down, we assume \( p^{(0)}_g \approx [P]K_{g,P} = 10^{-3} \). While these considerations set the magnitude of the parameters, functional considerations – such as the presence of both an “ON” and an “OFF” state in the core NANOG-OCT4-SOX2 network, or the ability to model a repressor – leads us to modify these values in certain cases.

### 3.3.1.3 Quasi Steady State Approximation

For the free transcription factor concentration \([H]\), we need to determine the amount of molecules bound to regulatory sites. Since a given factor may contribute to multiple complexes, we need to take into account each. If a certain complex \( A \) contains \( n_{A,H} \) molecules of specimen \( H \), and the concentration of this complex is \([A]\), then the total
amount of protein bound at regulatory sites is

$$[H_{\text{bound}}] = \sum_{g} \sum_{A \in g} n_{A,H}[A], \quad (3.14)$$

where the summation runs over all possible promoter sites $g$ and regulatory complexes $A$, that can form at a given promoter site as indicated by the symbolic summation rule $A \in g$. If a particular complex $A$, assembled at promoter $g$, consists of proteins $H, \ldots, I$, then according to Eq. (3.12)

$$[A] = [g]w_A, \quad (3.15)$$

where

$$w_A = \frac{[H]}{c_0} \cdots \frac{[I]}{c_0} \exp \left( \frac{E_A}{RT} \right). \quad (3.16)$$

Since the total promoter concentration of locus $g$ is $c^* \approx 1/\text{cell} \approx 0.1 \text{nM},$

$$c^* = [g] + \sum_{A \in g} [A] = [g] \left( 1 + \sum_{A \in g} w_A \right) = [g]Z_g, \quad (3.17)$$

thus,

$$[A] = \frac{c^*}{Z_g} w_A. \quad (3.18)$$

Using these notations, Eq. (3.3) can be expressed as

$$D_H = \frac{\partial[H_{\text{bound}}]}{\partial[H]} = \sum_{g} \sum_{A \in g} n_{A,H} \frac{\partial[A]}{\partial[H]} = c^* \sum_{g} \sum_{A \in g} n_{A,H} Z_g \left( \frac{\partial w_A}{\partial[H]} - \frac{w_A}{Z_g} \frac{\partial Z_g}{\partial[H]} \right). \quad (3.19)$$

Since $w_A$ is a polynomial of $H$,

$$\frac{\partial w_A}{\partial[H]} = n_{A,H} \frac{w_A}{[H]}, \quad (3.20)$$

and

$$\frac{\partial Z_g}{\partial[H]} = \sum_{B \in g} n_{B,H} \frac{w_B}{[H]}, \quad (3.21)$$

Introducing the sums

$$S_{g,H}^{(k)} = \sum_{B \in g} n_{B,H}^k \frac{w_B}{[H]}, \quad (3.22)$$
which are readily evaluated knowing the free specimen concentrations $[H]$, equation \((3.19)\) can be written as

$$D_H = c^* \sum_g \sum_{A \in g} \frac{n_{A,H}}{Z_g} \left( n_{A,H} \frac{w_A}{[H]} - \frac{w_A}{Z_g} \sum_B n_{B,H} \frac{w_B}{[H]} \right) =$$

$$= c^* \sum_g \left( \frac{S^{(2)}_{g,H}}{Z_g} - [H] \left( \frac{S^{(1)}_{g,H}}{Z_g} \right)^2 \right).$$ \hspace{1cm} (3.23)

If each complex $A$ can contain specimen H only once, then $S^{(2)}_{g,H} = S^{(1)}_{g,H}$.

If a transcription factor can also form a complex in addition to the one formed at gene regulatory sites, such as the dimerization of SOX2 and OCT4, then the QSSA needs to take that into account as well. In particular, assuming equilibrium, the dimer concentration $[OS]$ is given by

$$[OS] = K_{OS}[SOX2][OCT4].$$ \hspace{1cm} (3.24)

Thus $D_{OCT4}$ contains an additional term, $K_{OS}[SOX2]$, to those listed in Eq. \((3.23)\).

3.3.1.4 Model parameters

The magnitude of model parameters (Tables 3.1 - 3.3) were set by the following considerations. The transcription and translation rates were chosen in such a way that the steady state transcription factor (protein) concentrations are in the nanomolar range (in the order of 100 copies of the TF are present in the cell) when the promoter is fully active [83, 85]. To get a functional transcriptional regulatory system, the nanomolar concentration range must be also characteristic for promoter binding affinities, which by Eq. \((3.9)\) translates (at $T = 300$ K) into binding energies around 12 kCal/mol. The transcription factors were assumed to work through stabilizing RNAP II binding – with protein-protein binding energies around 4 kCal/mol [83]. This binding energy is increased for cooperative, multimolecular complexes. We assume that the probability of RNAP II binding in the absence of all the transcription factors considered is very low, $[P]K_{g,P} \approx 10^{-3}$. 
Chapter 3. Autocrine feedback in pluripotent stem cell maintenance

<table>
<thead>
<tr>
<th>gene</th>
<th>$B_{g,TF}$</th>
<th>C_{g,TF}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NANO</td>
<td>OCT4</td>
</tr>
<tr>
<td>Nanog</td>
<td>12</td>
<td>-99</td>
</tr>
<tr>
<td>Oct4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sox2</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Klf4</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Esrrb</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Fgf5</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Fgf4</td>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3.1. Model parameters I: binding energy and three-component cooperativity. $B_{g,TF}$ [kcal/mol]: binding energy of TF at locus $g$; $C_{g,TF}$ [kcal/mol]: cooperativity among RNAP II, TF and locus $g$.

<table>
<thead>
<tr>
<th>gene</th>
<th>$C_{g,TF_1,TF_2}$</th>
<th>$P$</th>
<th>$K_{g,P}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nanog</td>
<td>-99</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Oct4</td>
<td>0.02</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sox2</td>
<td>6</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Klf4</td>
<td>6</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Esrrb</td>
<td>6</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Fgf5</td>
<td>-99</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Fgf4</td>
<td>4</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2. Model Parameters II: four-component cooperativity and RNAP II binding. $C_{g,TF_1,TF_2}$ [kcal/mol]: cooperativity among RNAP II, TF1, TF2 and locus $g$, where $B_{g,TF_1}, B_{g,TF_2} \neq 0$; $K_{g,P}$ [1 nM]: binding probability of RNAP II (P) at locus $g$. The probability for RNAP II binding to the Oct4 locus was set to 0.001 in model variant A and B.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
<th>unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>translation and transcription rate</td>
<td>$\alpha_{gene}$</td>
<td>4</td>
</tr>
<tr>
<td>decay rates for TFs</td>
<td>$\delta_{TF}$</td>
<td>1</td>
</tr>
<tr>
<td>autocrine decay rates</td>
<td>$\delta_{FGF}$</td>
<td>10</td>
</tr>
<tr>
<td>OCT4/SOX2 association rate</td>
<td>$k_a$</td>
<td>1</td>
</tr>
<tr>
<td>OCT4/SOX2 dissociation rate</td>
<td>$k_d$</td>
<td>1</td>
</tr>
<tr>
<td>FGF receptor concentration</td>
<td>$R_{tot}$</td>
<td>1</td>
</tr>
<tr>
<td>receptor ligand binding energy</td>
<td>$B_{R,L}$</td>
<td>11</td>
</tr>
<tr>
<td>indirect inhibition rate</td>
<td>$a$</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 3.3. Model Parameters III.
3.3. Results

3.3.2 The core network

3.3.2.1 Core network scenarios

First, we consider various scenarios for the core NANOG circuit (Fig. 3.2) and compare their behavior. Steady state system behavior was characterized by numerically obtaining intersections of nullcline planes (Fig. 3.3).

Our starting point is the model A, which is symmetric in the roles of SOX2 and OCT4 (Fig. 3.2A). This model exhibits bistability: there are two stable fixed points corresponding to the “ON” and “OFF” states of the system, separated by an unstable fixed point. Linear stability analysis reveals that the stable fixed points are stable nodes, thus, no oscillations are expected in their vicinity.

Augmenting the model with a negative feedback through OCT4, as suggested by [68, 73], can be accomplished by increasing the binding affinity of the OCT4 protein to the Nanog regulatory site, and decreasing the stability of the OCT4-containing RNAP II complex (Fig. 3.2B). We assume that the binding affinity of OCT4 is lower than that of NANOG or the OCT4/SOX2 dimer – reflecting that high concentration of OCT4 (overexpression) was needed to elicit the inhibition. Once OCT4 is bound, however, we assume a strong inhibitory effect. As suggested in [73], this change indeed can transform the “ON” state from a stable node to a stable spiral, but only if the OCT4 binding affinity is higher than the values characteristic for the other TFs. In such a case the fluctuations in [OCT4] are of similar magnitude than that of [NANOG] (data not shown). As OCT4 levels appear quite stable in mouse embryonic stem cells ([76], Fig. 3.5), in models compatible with this observation the direct OCT4 negative feedback is unlikely to play an important role.

In model C (Fig. 3.2C), OCT4 is independent of NANOG activation. In such a scenario the same parameters that were used for model A yield only the “OFF” fixed point as Oct4 never turns on. Considering the steady presence of OCT4 in embryonic stem cells, we argue that for this model it is reasonable to choose an increased probability for RNAP II binding to the Oct4 locus even in the absence of SOX2. This choice yields a bistable system similar to that of model A.

Model D (Fig. 3.2D) is derived from model C by adding a NANOG autorepression feedback. As Fig. 3.3 demonstrates, this change does not alter substantially the systems dynamics as the Nanog promoter activity can be obtained by scaling the activity of model C (3.26).

Finally, we combine model D with model B to see if NANOG autorepression can further promote the transformation of the “ON” fixed point into a spiral. Adding
NANOG autorepression to model B reduces the equilibrium NANOG levels (as determined by the amount of OCT4 present) according to Eq. (3.26), which further stabilizes the fixed point. We also derive model E (Fig. 3.2E) by adding OCT4 as a Nanog repressor to model D. In this scenario we find the fixed point still strongly attractive: oscillations decay fast and change both NANOG and OCT4 levels to a similar extent.

This analysis of the core NANOG circuit variants suggest that they are likely to behave as bistable systems. Hence, the experimentally seen heterogeneity, given the stability of Oct4 expression levels ([76], Fig. 3.5), is an unlikely consequence of the core NANOG-OCT4-SOX2 dynamics. In the following we focus on model D, as the simplest variant of the investigated networks, that is functionally equivalent to the one proposed by the most recent experimental data [72]. This choice, however, is somewhat arbitrary as all model networks function as a bistable switch.

**Fig. 3.2. NANOG core circuit models studied in this work.** We consider the OCT4/SOX2 dimer as a common transcription factor for all three genes, and the NANOG protein to be a transcription enhancer for the Sox2 gene. We investigate model scenarios in which NANOG is (A) or is not (C) an inducer of Oct4. Furthermore, we consider further variants, such as the model proposed by Pan et al. where high OCT4 levels repress Nanog (B), as well as the NANOG autorepressor feedback proposed by [72] (D).
3.3. Results

3.3.2.2 Analysis of core networks

Steady state system behavior is characterized by numerically obtaining intersections of nullcline planes – steady state concentrations obtained when one of the specimen is kept at a fixed value. For example, by keeping [NANOG] at a predetermined value, the steady state concentration values satisfy Eq. (3.9) for both $g = Sox2$ and $Oct4$ so that the corresponding time derivatives are zero. Thus the obtained concentration values are at the intersection of the $d[SOX2]/dt = 0$ and $d[OCT4]/dt = 0$ nullclines.

Model A As the nullcline intersections demonstrate in Fig. 3.3A, steady state concentrations of [NANOG] is a monotonous increasing function of both [OCT4] and [SOX2] when either of these quantities are used as fixed control parameters. The curves [NANOG] vs [OCT4] and [NANOG] vs [SOX2] reflect the gradual activation of the Nanog promoter and the saturation of transcription at transcription factor (TF) concentrations exceeding the promoter’s binding affinity. The markedly nonlinear functional form is due to the need of TF dimerization to activate transcription. Fig. 3.3A also demonstrates that steady state [OCT4] and [SOX2] levels decrease for large concentrations of the complementary species as formation of the OCT4/SOX2 complex depletes the pool of free proteins. This simple model exhibits bistability: there are two stable fixed points corresponding to the “ON” and “OFF” states of the system, separated by an unstable fixed point. Linear stability analysis reveals that the stable fixed points are stable nodes, thus, no oscillations are expected in their vicinity.

Model B A negative feedback through OCT4, as suggested by [68, 73], can be accomplished by increasing the binding affinity of the OCT4 protein to the Nanog regulatory site, and decreasing the stability of the OCT4-containing RNAP II complex. Fig. 3.3B reveals that such an inhibition renders the [NANOG] vs [OCT4] curve decreasing for OCT4 concentrations higher than its binding affinity. As suggested in [73], this change indeed can transform the “ON” state from a stable node to a stable spiral, but only if the OCT4 binding affinity is higher than the values characteristic for the other TFs. In such a case the fluctuations in [OCT4] are of similar magnitude than that of [NANOG] (data not shown).

Model C To reflect the increased probability for RNAP II binding to the Oct4 locus, in Fig. 3.3C for the Oct4 promoter we use the value $[P]K_{Oct4,P} = 0.02$ instead of the value 0.001 – used for all other promoters. With this choice the intersection of the OCT4 and SOX2 nullclines moves to higher OCT4 values, and yields a bistable system where two stable nodes are separated by an unstable fixed point. Since in this
model OCT4 production only depends on the dimer concentration, [OCT4/SOX2], the [OCT4] vs [SOX2] curve is the same irrespective if we set [SOX2] directly or indirectly through [NANOG].

**Model D** As Fig. 3.3D demonstrates, a NANOG autorepression feedback does not alter substantially the systems dynamics: when Nanog is turned on, the steady state NANOG concentrations are somewhat reduced compared to the values of model C (gray lines). Potent autorepression does not change $Z_{\text{ON}}^{\text{Nanog}}$ as

$$\exp(C_{\text{Nanog},\text{NANOG}}) = \exp(C_{\text{Nanog},\text{NANOG,OS}}) \approx 0. \quad (3.25)$$

In contrast, autorepression increases $Z_{\text{OFF}}^{\text{Nanog}}$ by $[\text{NANOG}]K_{\text{Nanog, NANOG}}$. The probability of transcription in the models with or without autorepression, $p'_{\text{Nanog}}$ and $p_{\text{Nanog}}$, respectively, can be directly compared by a suitable scaling

$$p'_{\text{Nanog}} \approx p_{\text{Nanog}} \frac{1}{1 + [\text{NANOG}]K_{\text{Nanog, NANOG}}Z_{\text{Nanog}}}. \quad (3.26)$$

**Model E** Adding OCT4 as a Nanog repressor to model D renders the [NANOG] vs [OCT4] curve decreasing for OCT4 concentrations higher than its binding affinity. The fixed point remains strongly attractive: oscillations decay fast and change both NANOG and OCT4 levels to a similar extent.

### 3.3.3 Sensitivity of the core network to model parameters

Changes in model parameter values can gradually shift the nullclines and fixed points in the phase space. As changes in nullcline positions and shapes can create or remove intersections, the presence of both the “OFF” and “ON” states are parameter dependent.

To gauge the model’s sensitivity to parameter values, we systematically vary all of them, one-by-one, by 20% and 40%. Starting the simulations from the “ON” fixed point, we obtain the new steady state values under the altered parameter setting. As Fig. 3.4 demonstrates, most parameters effect only one molecular species directly, and the fixed point moves along a corresponding nullcline intersection line. In particular, if the “ON” fixed point falls onto the saturated regime of the nullcline intersections, then the steady state concentrations of model components, such as [NANOG] and [OCT4] may be differently altered. For example, changing the binding energy between RNAP II and the OCT4/SOX2 dimer at the Nanog promoter by 20% shifts the
3.3. Results

Fig. 3.3. Steady state properties of the NANOG core circuit model variants shown in Fig. 3.2. Each panel is annotated as the corresponding model variant (A-E). Steady state concentrations were obtained in simulations where one of the molecular specimen was kept at a fixed value (NANOG: orange, OCT4: green and SOX2: blue). Intersection of all three curves indicate fixed points. Two of these fixed points, located at low and high concentrations are stable nodes in models A, C and D: a perturbed system is expected to return to these states without oscillations. In model B and E the fixed point at high concentrations may turn into a stable spiral. In contrast, the fixed point at intermediate concentration is unstable. Thus, the system behaves as a bistable switch, with distinct “ON” and “OFF” states. To ease comparison between model variants, nullclines of models A, C and D are plotted with gray lines in panels B, D and E, respectively. Concentrations are presented in nM units.
Fig. 3.4. Parameter sensitivity analysis of the “ON” fixed point of the core NANOG circuit. Each model parameter was changed by 20% and 40% and the obtained new steady states (gray circles) are overlaid on the phase space plot of Fig. 3.3D. Of particular interest is a change in $C_{\text{Nanog,OS}}$, the increase of binding energy of RNAP II at the Nanog promoter in the presence of the OCT4/SOX2 dimer. Twenty percent change in this parameter shifts the equilibrium NANOG concentration by more than an order of magnitude more than that of OCT4 (black circle). B): Parameter dependence of nullclines and fixed points: the thick curves were obtained in a simulation with increased $C_{\text{Nanog,OS}}$. Notice that in the “ON” state, NANOG concentration is increased by a factor of 2, while OCT4 levels remained the same. Furthermore, the altered set of parameters excludes the “OFF” state of the system. Concentrations are presented in the units of nM.

equilibrium NANOG concentration by more than an order of magnitude more than that of OCT4. This observation is the basis of our explanation for NANOG heterogeneity and OCT4 homogeneity within a cell population.

### 3.3.4 NANOG heterogeneity

As we demonstrated above, the most plausible assumptions do not suggest the presence of substantial oscillations within the core NANOG-OCT4-SOX2 system. To explain the observed broad distribution of NANOG expression within a population of mouse ES cells, we hypothesize that model parameters such as binding energies depend on the larger biochemical context. For example, ERK activity is a known potent negative regulator of Nanog transcription [75, 76]. In turn, FGF signaling is capable to create an autocrine feedback using cell surface receptors that feed into the MAPK pathway, and FGF activity is indeed a well established modulator of ES cell heterogeneity [86]. Nodal signaling gives an alternative possibility for an autocrine extracellular regulation of Nanog, mediated through the Smad family [76].
Fig. 3.5. The behavior of various signaling components during NANO2 fluctuations. A): Expression levels, as determined by rt-pcr, of NANO2, SOX2, OCT4, ESRRB and the FGFs active within mouse ESCs. Data show fold change differences in expression normalized to Gapdh transcript levels. Each expression value is the average of values obtained from three independent experiments. B): ERK activity in cell populations with various extent of NANO2 expression. As a positive control, we also include pERK western blot data from mouse embryonic fibroblasts (MEF). C): Western analysis of dox-inducible NANO2 ES cell line. Note reduction of pERK levels in response to increasing amounts of NANO2.

To demonstrate that autocrine signaling loops can induce fluctuating NANO2 expression levels, we consider a feedback through FGFs and ERK. As a particular example, we investigate a scenario in which the stability of the RNAP II complex containing the OCT4/SOX2 dimer depends on ERK activation. To explore the behavior of the FGF pathway during NANO2 fluctuations, mouse ES cells were sorted based on their NANO2 expression levels into four groups (low, medium, high, very high). For each group we determined the transcriptional activity of key genes, such as Oct4, Esrrb, and FGF family members expressed by mouse ES cells: Fgf4, Fgf5 and Fgf8. As Fig. 3.5 demonstrates, the range of Oct4 variability is less than half of that of NANO2. In contrast, the range of variability in the expression levels of Fgf4 and Esrrb is even greater than that of NANO2. High Fgf4 expression is associated with high NANO2 expression, whereas high Fgf5 and Fgf8 expression is characteristic for cells with low levels of NANO2. ERK activity (phosphorylation) assays revealed an inverse relationship between ERK activity and NANO2 expression levels (Fig. 3.5). This correlation between NANO2 expression and ERK activity is seen in both sorted NANO2 subpopulations (Fig. 3.5B) and in a doxycycline-inducible NANO2 ES cell line (Fig. 3.5C).
Fig. 3.6. The signaling and transcriptional network considered to regulate mouse ES cell maintenance. A): Full model. Green, red and blue arrows represent transcriptional activation, repression and translation, respectively. Black arrows represent complex formation, and various multi-step processes: (i) FGF secretion resulting in an effective autocrine ligand concentration and (ii) activation of the MAPK. B): Schematic representation of the autocrine feedback loop. The suggested model acts as a noisy negative feedback regulator which includes a signal amplifier and delay.

Based on known regulatory binding sites [87], and the expression data in Fig. 3.5, we consider the autocrine feedback loop shown in Fig. 3.6. Using the available transcription factor-DNA binding ChiP data set, we selected two transcriptional regulators for each gene. In particular, we assume that the KLF4-ESRRB system is downstream of the NANOG core circuit. We suggest that NANOG, KLF4 and ESRRB are activators and repressors of the two Fgf genes considered in the model. Finally, we assume that secreted FGF proteins bind to cell surface receptors in an autocrine manner. The activation of FGF receptors initiate the intracellular MAPK signaling pathway [88], which closes the feedback to Nanog through modulating the binding affinities of the OCT4/SOX2 dimer, the only transcription factor of NANOG explicitly considered in the model. In the following we build up this complex signaling model from simpler modules.
3.3. Results

Fig. 3.7. The KLF4-ESRRB module as an amplifier. We consider an autocatalytic regulation of KLF4, and cooperative positive regulation from NANOG (A). Keeping [NANOG] at various predetermined values, we obtained the steady state concentrations and promoter activity of Klf4 and Esrrb. The relationship between these values and [NANOG] is strongly non-linear, reflecting the autoregulation and cooperative binding with NANOG. The solid line represent fitted Hill curves with $n = 2.3 \pm 0.2$ and $n = 6.2 \pm 0.4$ for KLF4 and ESRRB, respectively (B, C). Concentrations are presented in the units of nM.

3.3.5 KLF4 and ESRRB

Klf4 and Esrrb are known to be pluripotency genes and as a recent study exposed, both are direct targets of NANOG [89]. KLF4 was reported to bind to it’s respective promoter as well as to the promoter regions of Nanog and Esrrb [90]. To keep our model as simple as possible, we restrict the number of regulatory connections to two per locus. We further assume that Esrrb is more downstream than Klf4 is (see Fig. 3.7A). Promoter binding affinities were chosen in such a way that [NANOG] is in the nanomolar range when Klf4 and Esrrb genes switch on. With such assumptions, the NANOG-KLF4-ESRRB cascade can function as an amplifier. By keeping [NANOG] steady (as input), the autocatalytic expression levels of Klf4 are well approximated by a Hill function of exponent 2 (Fig. 3.7B). The steady state expression level of Esrrb is an even more non-linear function of [NANOG]: the abrupt switch is steeper than a Hill function with $n = 6$ (Fig. 3.7C). Thus, consistent with empirical data, if the KLF4-ESRRB system is tuned in this regime of operation, changes in the transcriptional activity of Esrrb may exceed by an order of magnitude that of Nanog.

3.3.6 Autocrine FGF signaling

Chromatin immunoprecipitation assays revealed that the Fgf4 gene has binding sites for KLF4 and ESRRB, and the Fgf5 gene has sites for NANOG and ESRRB [91, 92]. Microarray expression analysis [81] has shown that FGF Receptor 1, that can bind all three flavors of secreted FGFs, is expressed by mouse ES cells. Thus, we work with
an aggregate autocrine FGF concentration, \([\text{FGF}^*]\) to determine downstream receptor activity. As \(\text{Fgf5}\) and \(\text{Fgf8}\) expression levels appear to be similarly regulated, in our model both are represented by \(\text{Fgf5}\). While the dynamics of autocrine FGF signaling has not been studied in mouse embryonic stem cells, autocrine EGF signaling was explored extensively in other experimental systems [31]. Experiments with autocrine EGF signals revealed a linear relationship between cell surface autocrine ligand concentration and the production rate of the protein. Thus, we assume that the equation governing autocrine FGF ligand concentration, \([\text{FGF}^*]\), is

\[
\frac{d[\text{FGF}^*]}{dt} = \sum_{i \in \{4,5\}} \frac{\alpha_{\text{Fgf-i}} p_{\text{Fgf-i}} - \delta_{\text{FGF-i}}[\text{FGF-i}]}{1 + D_{\text{FGF-i}}},
\]

(3.27)

where \(0 < p_{\text{Fgf-i}} < 1\) is the probability of transcription at the \(\text{Fgf4}\) and the \(\text{Fgf5}\) locus, and \(\delta_{\text{FGF-i}}\) is the decay rate combined with the diffusive flux transporting the ligand off the cell surface. The \(\alpha_{\text{Fgf-i}}\) coefficient reflects both production and the conversion between autocrine ligand concentration and production. In the EGF system the relationship between steady state autocrine ligand concentration and its production rate was \(0.05 \text{ pM} / (\text{molecules/cell/h}) \approx 0.05 \text{ pM} / (30 \text{ pM/h}) \approx 0.1\% \text{ h}\) (see Fig. 6 of [31]). Based on this result, we expect that less than 1% of the FGF molecules produced in an hour (our approximate time unit) will act as autocrine ligands at the cell surface. Therefore, providing autocrine ligand concentration in the nanomolar range requires higher production rates than the rates we assumed for transcription factors (see Table 3.1).

The MAPK cascade, the signaling pathway downstream of the FGF receptor, has been studied extensively both by computational and biochemical methods [2, 93, 94]. These studies revealed two characteristic operation mode: changes in receptor ligation may elicit a transient and a sustained ERK activity. Interestingly, both responses are well approximated by a linear response function. Since the characteristic lifetime of the transient response is in the order of minutes, we assume that the relatively slow changes in \(\text{Nanog}\) expression that take place over several hours reflect a sustained change in steady state ERK activation. Our experimental data (Fig. 3.5) is also consistent with a change in the steady state ERK activity. Hence, we propose that the normalized difference

\[
\epsilon = \frac{[\text{ERK}^*] - [\text{ERK}^*]_0}{[\text{ERK}^*]_0}
\]

(3.28)

between ERK activity (i.e., concentration of phospho-ERK, \([\text{ERK}^*]\)) and a reference
(baseline) level activity \([\text{ERK}^*]_0\) is proportional to the number of active receptor complexes \(R^*\) as

\[
\varepsilon \sim R^*.
\]  

(3.29)

The steady state concentration of active receptors is

\[
R^* = R_{\text{tot}} \frac{[\text{FGF}^*]}{1/K + [\text{FGF}^*]},
\]  

(3.30)

where \(R_{\text{tot}}\) is the total amount of FGF receptors at the cell surface, and \(K\) is the binding constant between the receptors and their FGF ligands.

The connection between ERK activity and \(\text{Nanog}\) activation is currently unknown. Here we assume, that the regulation involves the modulation of the binding affinity of the OCT4/SOX2 dimer, the only transcription factor of \(\text{Nanog}\) that is explicitly considered in the model:

\[
C_{\text{Nanog,OS}} - C_{\text{Nanog,OS}}^{(0)} \sim \varepsilon,
\]  

(3.31)

thus

\[
C_{\text{Nanog,OS}} - C_{\text{Nanog,OS}}^{(0)} = aR_{\text{tot}} \frac{[\text{FGF}^*]}{1/K + [\text{FGF}^*]},
\]  

(3.32)

where the coupling factor \(a\) is chosen in such a way that for a typical simulation the magnitude of the right side of Eq. (3.32) is smaller than one.

### 3.3.7 NANOG expression determined by autocrine feedback

To explore the behavior of the full model (Fig. 3.6), we first consider steady states obtained with various (fixed) values of \(C_{\text{Nanog,OS}}\), the binding energy between RNAP II and the OCT4/SOX2 dimer at the \(\text{Nanog}\) promoter – the assumed site of ERK regulation (Fig. 3.8). When \(C_{\text{Nanog,OS}}\) is set by the feedback Eq. (3.32) with a specific “gain” parameter \(a\), the system reaches a single steady state that falls on the steady state curves shown in Fig. 3.8. We propose that the observed NANOG heterogeneity is resulted by slow alterations in model parameters – specific for individual cells – like the feedback strength or the efficiency of autocrine ligand capture. For strong enough autocrine feedbacks the system is characterized by a fold in the phase space (pitchfork bifurcation), hence small changes in the parameter values can have disproportionally large effects on the steady state concentration values. Furthermore, due to the hysteresis distinct substates can co-exist (Fig. 3.9), hence the concentration values also reflect the history of the system. All these substates are, however, still within the “ON” state.
Fig. 3.8. The effects of modulating the transcriptional regulation of Nanog. Steady state protein concentrations are plotted for various values of $C_{\text{Nanog,OS}}$, the binding energy among RNAP II, the OCT4/SOX2 dimer and the Nanog locus. If the binding energy is set by ERK activity through the feedback (Eq. 3.32), the system evolves into a steady state (gray symbols) which depends on parameter $a$, the strength of the feedback.

Fig. 3.9. Presence of substates within the “ON” state of the core Nanog switch.
3.3. Results

Fig. 3.9. Presence of substates within the “ON” state of the core Nanog switch. (cont.)

Steady state protein concentrations obtained for strong autocrine feedback ($a = 2$) as a function of the autocrine ligand decay parameter $\delta_{FGF}$ which is sensitive to the efficiency of autocrine ligand capture (A). The abrupt change and the hysteresis indicates the existence of distinct substates. Both substates are within the “ON” state of the core NANOG circuit as the parameters identified in Fig. 3.4, like $B_{Sox2,NANOG}$, the binding energy of NANOG to the Sox2 promoter, are still able to shut off Nanog expression (solid symbols).

The switch between the “ON” and “OFF” states continues to involve a bifurcation as the sudden jumps and hysteresis indicates (B). The “ON” and “OFF” states as well as the two substates can be visualized in a three dimensional parameter space, where the steady state NANOG concentration is plotted as a function of the autocrine ligand decay parameter $\delta_{FGF}$, and the binding energy of NANOG at the Sox2 locus, $B_{Sox2,NANOG}$ (C). Concentrations are presented in the units of nM.
of the NANOG core circuit: the parameters identified in Fig. 3.4, like the binding affinity of NANOG to the Sox2 promoter, are still able to shut off Nanog expression through another sudden change (Fig. 3.9C). To address the robustness of the substates, we performed a systematic variation of model components by 10%. We found, that when both distinct “ON” and “OFF” states were present, the “ON” state exhibited substates separated by a bistable region in 85% of the cases.

Large fluctuations readily develop as a response to a high frequency noise added to the parameters. As an example, a 30% modulation of the FGF decay parameter $\delta_{FGF}$ results in slow, but large amplitude transitions between the substates (Fig. 3.10A-C): the model-predicted duration of a transition is in the order of a day, and the typical time of the system spends in the same substate is in the order of a week. Such simulations also allow to correlate expression levels in the model with experimental data shown in Fig. 3.5. In the time series we identified regimes where [NANOG] was below 0.7 nM (“NANOG low”) or above 1.0 nM (“NANOG high”). For both types of time intervals we averaged the expression level of each factor, and normalized it to the “NANOG high” state (Fig. 3.10D). The general tendencies of both Fig. 3.10D and Fig. 3.5 are identical: Nanog expression levels are more variable than that of Oct4, and the KLF4-ESRRB amplifier can further increase the variability in Fgf expression.

### 3.4 Discussion

#### 3.4.1 The mechanism behind Nanog fluctuations

Protein and mRNA levels can fluctuate due to the inherently stochastic kinetics of gene transcription and translation. Even an unregulated, constitutively expressed gene exhibits a 30% spread of expression levels over a population of identical cells [95]. Clearly, the reported variations in the stem cell maintenance network, especially those genes that are downstream of Nanog, greatly exceed this baseline variability. Thus, we expect that the observed heterogeneity in Nanog expression levels is generated by the dynamics of the regulatory system as it greatly amplifies the molecular stochastic noise. To describe a potential mechanism generating the experimentally observed dynamics, we proposed a model that operates with binding affinities of multiple transcriptional regulators of pluripotency, the topology of the transcriptional regulatory networks and activity of an autocrine signaling pathway. The resulting model produces varied expression levels of several components of pluripotency regulation, largely consistent with our (Fig. 3.5) and previously reported empirical observations [76].
3.4. Discussion

Fluctuations in protein concentrations (A) and gene expression (C), driven by a noise ($n_\delta(t)$) (B) added to the autocrine ligand decay parameter $\delta_{\text{FGF}}$, which changes its value by $\pm 30\%$. In the time series regimes with high and low NANOG concentration were identified, using threshold concentrations of 1.0 nM and 0.7 nM for Nanog high and low expression, respectively. The corresponding time-averaged expression levels are plotted in panel (D).

Fig. 3.10. Fluctuations in protein concentrations and gene expression driven by a noise. Fluctuations in protein concentrations (A) and gene expression (C), driven by a noise ($n_\delta(t)$) (B) added to the autocrine ligand decay parameter $\delta_{\text{FGF}}$, which changes its value by $\pm 30\%$. In the time series regimes with high and low NANOG concentration were identified, using threshold concentrations of 1.0 nM and 0.7 nM for Nanog high and low expression, respectively. The corresponding time-averaged expression levels are plotted in panel (D).
We argue that slow fluctuations in *Nanog* expression likely reflect individual cell-specific changes in the parameters of an autocrine feedback loop, such as changes in ligand capture efficiency, receptor numbers or the presence of cross-talks within the MAPK signal transduction pathway. While high-frequency variability may be filtered out by the slow dynamics of transcription factor synthesis and accumulation, low frequency changes, such as a slow alteration in the cell’s microenvironment, are capable to push expression levels across substate boundaries. Given the complexities of a cell’s variable exposure to autocrine/juxtacrine signaling in culture, this model incorporates a plausible basis for a variable activity of an autocrine signaling pathway eliciting heterogeneous expression of intracellular components. In this view the fluctuations of *Nanog* reflect the response of a regulatory system with multiple feedbacks in a non-stationary environment.

This feedback mechanism, that does not involve changes in OCT4, is consistent with both the stability of Oct4 expression levels [76], (Fig. 3.5), and the observation that alteration of OCT4 levels induces ES cells to differentiate rapidly; too much or too little OCT4 rapidly directs cells to differentiate [96, 97]. Very small increases in Oct4 expression causes differentiation to mesoderm and endoderm, and reduction of Oct4 levels induces loss of pluripotency and dedifferentiation to trophectoderm.

### 3.4.2 Comparison with experimental data

The proposed computational model needs to be compared with experimental data on at least three levels. (1) The dynamics of the whole system can be evaluated and compared with corresponding experimental data. (2) Our model makes explicit or implicit assumptions on the topology of the transcriptional regulatory network, and identifies functional modules. Finally, (3) most model parameters are expressed either as molecular binding affinities or as parameters effecting the stability of the transcription complex in addition to production and decay rates.

#### 3.4.2.1 Dynamics

Our model calculations demonstrate that a biologically plausible autocrine feedback can create distinct substates within the “ON” state of the core NANOG switch. While stochasticity and feedback regulation has been proposed to explain Nanog fluctuations [73, 85, 98], the previously proposed mechanisms involved noise induced transitions between the “ON” and “OFF” states of the core Nanog switch. Experimental evidence, however, suggests that Nanog expression in Nanog-low ES cells is still much higher than that in cells committed to differentiation [76]. Furthermore, Nanog-low cells can
still be maintained indefinitely without committing to differentiation – their Nanog expression level will, in fact, increase. These empirical observations clearly support a mechanism that operates with transitions between substates that do not involve the “OFF” state of the core Nanog switch. The existence of distinct substates of Nanog expression was also suggested by a recent study, that analyzed experimentally observed changes in Nanog expression profiles in terms of a phenomenological model, that made no explicit assumptions on the underlying molecular signaling network [77].

Our approach also demonstrates that the slow modulation of Nanog expression can reflect changes that are external to the core circuit – instead of the stochastic expression of the transcription factors [85]. The topology of the regulatory network considered here is more elaborated than in previous studies [73, 85], and thus able to represent and predict changes in expression of several downstream genes. Furthermore, our model does not include direct autoregulation of Nanog – a frequent assumption in simplified models which lacks empirical support.

FGF and NODAL signaling are clearly active in mouse embryonic stem cells, and function in an autocrine fashion in undifferentiated cells [86]. Previous studies [75, 76] have shown that these autocrine signaling pathways influence the dynamic heterogeneity of ES cells in culture, likely through specific molecular mechanisms that have yet to be elaborated. Our results show highly variable expression of FGF ligands and intracellular ERK activity in ES cells grown in serum-based media (Fig. 3.5). The model behavior presented in Figs. 3.10 is largely consistent with the anticorrelation found in ERK activity and Nanog expression (Fig. 3.5) as well as existing information on FGF signaling [75].

3.4.2.2 Regulatory network topology

The regulatory network shown in Fig. 3.6 contains motifs that are well established as well as hypothetical regulatory connections that we explore in this work. The full network can be broken up into four modules: the core NANOG switch, the downstream ESRRB amplifier, the autocrine FGF module and finally, the feedback through the MAPK cascade.

Our analysis indicates that each of the recently proposed regulatory architecture of the core NANOG-OCT4-SOX2 network (Fig. 3.2) functions as a bistable switch. We argue, that irrespective of the underlying model details, the nonlinearity needed to create two stable fixed points is provided by the autoregulation and the dimerization of the OCT4/SOX2 transcription factors. The switch-like behavior is also maintained when OCT4 acts as a likely low affinity repressor – an assumption motivated by the
finding that OCT4 overexpression was required for the manifestation of the repressive behavior [99].

The roles KLF4 and ESRRB play in stem cell maintenance are in the focus of recent scientific interest [89, 100, 101]. Based on chromatin immunoprecipitation data [87], both the \textit{Klf4} and \textit{Esrrb} genes have several potential regulatory sites, which include transcription factors from the NANOG core as well as allow autoregulatory feedbacks. We demonstrated that a subset of the known likely regulatory interactions is capable to function as an amplifier, greatly expanding the variability of these factors. While the existence of (functional) autoregulation in these loci has not yet been established by targeted experiments, the proposed amplifier function of the KLF4-ESRRB system is in good agreement with the experimentally obtained large variations in \textit{Esrrb} (Fig. 3.5 and [102]) and downstream \textit{Fgf} expression levels (Fig. 3.5).

The \textit{Fgf} genes which are known to be actively transcribed in mouse ES cells exhibit regulatory sites for the transcription factors considered in this work. Chromatin immunoprecipitation data reveals that \textit{Fgf4} has putative regulatory sites for KLF4 and ESRRB, \textit{Fgf5} has putative regulatory sites for NANOG and ESRRB, while \textit{Fgf8} exhibits a regulatory site for NANOG [87]. While the nature of the regulatory connection is currently unknown, based on the strong correlation between \textit{Nanog} and \textit{Fgf4} expression and the strong anticorrelation between \textit{Nanog} and \textit{Fgf5} as well as \textit{Nanog} and \textit{Fgf8} expression levels, we assume that \textit{Fgf5} and \textit{Fgf8} are repressed while \textit{Fgf4} expression is enhanced by the transcription factors considered in this work. The repressional regulation of \textit{Fgf} genes is of key importance – it yields a negative feedback loop, and this will be tested by targeted experiments in the future.

Finally, the MAPK module and the autocrine ligand concentration was treated phenomenologically by simplified input-output relationships reflecting previously established results. In particular, we postulated a linear relationship between autocrine ligand concentration and its rate of production [29, 31]. The MAPK pathway can exhibit transient and sustained activation [103, 104], the latter lasting only for less than an hour [94]. As changes in transcriptional regulation are operating on a slower time scale, here we considered a change in sustained ERK activation, which was also carefully studied and is expected to operate as an amplifier (ultrasensitive switch) linking input and output with a Hill coefficient much greater than one [93, 105]. This observation motivates our use of a coupling factor \( a \) in Eq. (3.32) that is able to substantially reduce \textit{Nanog} expression, in agreement with specific observations in mouse ES cells [75, 76].
3.5. Conclusions

3.4.2.3 Parameter values

A molecular regulatory network model is bound to have several parameters, for most of which very little empirical data is available. As one way to contain this problem, here we used a modular approach: first we identified the behavior of smaller units, like the core NANOG switch, or the ESRRB amplifier. Functional requirements can considerably constrain the possible parameter values within such a module. As a second effort to reduce the arbitrariness of model parameters we derived expression rates from binding energies[82, 83]. Finally, unless empirical evidence suggested otherwise, we kept the simplest (uniformly assigned) values of the parameters and did not try to match empirical data by fine tuning the parameters. Nevertheless, we do not consider our parameters to be predictive – many different combinations can give similar overall behavior. Yet, they are important to demonstrate that the proposed regulatory system can work with plausible parameter values.

3.4.3 Future directions

The major focus of this model involves a regulatory network of NANOG, OCT4, and SOX2; these transcriptional regulators constitute a semi-independent regulatory module (“core” module [106]) in maintaining pluripotency. Clearly many other transcriptional regulators are involved in this complex module however. Our model represents an extensible platform for adding further components of the core module. The integration of other transcriptional regulators such as TCF3 [107], TBX3 [108], and signaling pathways such as Nodal and BMP signaling [76], will be an important future application of this model to study the inherent instability of the pluripotent phenotype in serum-based media. The model presented here does not take into account potential role of allelic expression of the Nanog locus [109] in directing heterogeneity in Nanog expression. Experimental data for this phenomenon are contradictory, however [110]. Further experimental data will be required to confirm this phenomenon, and if so, how this unusual mode of transcriptional regulation may be integrated into the model presented here.

3.5 Conclusions

The pluripotency of embryonic stem cells is sustained by a core cluster of co-regulated transcription factors, NANOG, OCT4 and SOX2. Surprisingly, NANOG as well as several other downstream transcription factors exhibit widely fluctuating expression
in a population of undifferentiated stem cells. To explain the observed heterogeneity of expression levels, we propose a computational model that couples the transcriptional regulation of Nanog to autocrine extracellular signals. We argue that the likely source of fluctuations is not the core regulatory cluster, but stochasticity within the autocrine feedback loops. The model predicts fluctuating expression levels for several factors involved in pluripotency maintenance, largely consistent with empirical observations presented here or reported previously. Our model indicates the presence of distinct substates of pluripotency, each exhibiting various expression levels of Nanog and downstream transcription or signaling factors. Thus, we predict that subpopulations within undifferentiated embryonic stem cells can have non-uniform responses to extracellular stimuli. Finally, we assign an empirically testable function to the transcriptional regulators KLF4 and ESRRB, and predict differential regulation of FGF family members. In vivo, we expect autocrine feedbacks to be highly sensitive to alterations of the stem cell microenvironment, hence our model can guide future studies linking stem cell maintenance to physical and biochemical properties of the stem cell niche.
Chapter 4

Role of soluble VEGFR1 signaling in vascular patterning

4.1 Background

Diffusion-limited growth can create various branching patterns, from snowflakes to bacterial colonies [111]. While these processes are highly diverse in physical details, each share the following common mechanism. A moving boundary, for example the surface of a crystal, interacts with a diffusive field, like the temperature. The dynamics of the field and the boundary is mutually interdependent as the expansion alters the surrounding field, while the field at the boundary determines the speed of expansion. This interdependent dynamics can give rise to the Mullins-Sekerka instability: steep gradients develop in the field around a protruding tip and promote its further extension. In this process, branching structures can be generated by progressive splitting of growing tips.

Vascular networks are also established by a series of branching events. Each vascular segment extends autonomously and contains several endothelial cells. While vascular patterning likely utilizes a variety of guidance mechanisms [52, 112, 113], one of the best established regulator of vascular growth is the response of endothelial cells to growth factors within the tissue environment, in particular to vascular endothelial growth factor type A (VEGF) [114–118]. In cultured endothelial cells, VEGF induces cell motility [119, 120], proliferation [121, 122] and directed movement (chemotaxis) [121, 123]. The chemotactic response is a sustained bias towards the growth factor gradient in the otherwise randomly directed cell movements [124]. While growth factors readily diffuse in culture, in a tissue environment growth factors are sequestered within the extracellular matrix, restricting both their diffusion and availability [125, 126].
Recent experiments demonstrated that secreted type 1 VEGF receptors (sVEGFR1 or sFlt1) can modulate vascular patterns. In particular, biochemical data indicates that sVEGFR1 can function as a “decoy receptor” by binding and inactivating VEGF within the tissue microenvironment [127–130]. While the absence of sVEGFR1 does not fully eliminate vascular sprouts within mouse embryonic stem cell-derived embryoid bodies, the direction of and distance between the sprouts becomes irregular when endothelial cells are deficient in expressing sVEGFR1. Re-introduction of sVEGFR1 expression into deficient cells restored normal vascular morphology [129, 131]. Furthermore, vascular sprouts elongate faster when endothelial cells immediately adjacent to the sprouts express more sVEGFR1. In summary, the following patterning mechanism is consistent with the available experimental data: when endothelial cells secrete sVEGFR1, the diffusing decoy receptors antagonize the pre-existing VEGF within the tissue microenvironment. Interestingly, in this guidance mechanism the concentration of active (not antagonized) VEGF forms a gradient pointing away from endothelial cells. This patterning mechanism therefore operates with a functional VEGF gradient that is the opposite of what was predicted by previous models aimed to explain vascular patterning [132–134].

While the biological foundation of sVEGFR1-related vascular guidance is well explored, less is known how these processes modulate vascular patterns. Expansion of the cells is a simple consequence of cells being programmed to follow an outward directed gradient. However, could such a mechanism in itself promote sprouting (self-organized branching), and how does the pattern change when parameters such as the lifetime or the affinity of the diffusive inhibitor changes? To understand the connection between the multicellular-scale organization and the molecular signaling mechanism, we investigated computational models of the core patterning process. In particular, we considered sVEGFR1 to be a diffusive inhibitor of VEGF, which promotes the expansion of the vasculature. While endothelial cells secrete sVEGFR1, most of the VEGF is sequestered in the ECM environment.

To represent a biological system, a typical mathematical model makes several – often implicit – assumptions. Most of these modeling choices are thought to be irrelevant and not driving the behavior emerging within the model. To demarcate the relevant and irrelevant model details, one can use multiple complementary modeling approaches: the same biological mechanism, thought to be relevant, can be represented by distinct models that can differ greatly in several modeling choices. When the complementary models yield the same behavior, the particular hidden or implicit assumptions in each model are thus likely irrelevant. We explore if and when a diffusive inhibitor can generate branching patterns. We introduce two, complimentary
4.2. Materials and methods

4.2.1 Cell culture

Human umbilical vein endothelial cells (HUVEC, Lonza) were maintained in EGM-2 medium (Lonza) under normal cell culture conditions: 37°C with a humidified 5% CO₂ atmosphere. HUVEC aggregates and sprouts were studied in an assay medium prepared by supplementing EGM-2 with 80 nM PMA (Merck), 40 ng/ml bFGF (Pierce), 40 ng/ml VEGF-A₁₆₅ (Pierce) and 50 µg/ml ascorbic acid (Sigma). Cells were fixed with 4% paraformaldehyde for 30 minutes at room temperature. Fixed cultures were stained with 0.25 mg/ml toluidine blue solution in phosphate buffered saline (PBS, pH 7.4).

4.2.2 VEGFR1 inhibitor

Anti-VEGFR1 antibody [AP-MAB0702, Abcam] was used to sequester soluble isoform of VEGFR1. This antibody is directed against the extracellular domain of recombinant human VEGFR1. Based on earlier data [135], 20 µg/ml of final concentration was used in the assay medium to sequester the soluble receptors.

4.2.3 Spheroid formation

Cell aggregation chambers were made by casting liquid 2% agarose (Invitrogen) in a PDMS micromold (3D Petri Dish, Microtissues) and allowing it to gelate. The micromold contains 35 pillars that define 35 wells with a diameter and depth of 800 µm.
Chapter 4. Role of soluble VEGFR1 signaling in vascular patterning

After removing the agarose gel from the micromold and placing it into a 35 mm culture dish (Greiner), the wells were equilibrated with cell culture medium for 2 hours before transferring the cells.

Cell suspensions were obtained by incubation of the cell monolayer with trypsin-EDTA (Sigma) for 2 minutes and washing the cells off. Cell suspensions were transferred into the wells formed by non-adherent agarose walls. An average of 3000 cells were transferred into each well and the agarose chamber was filled up with EGM-2 medium and kept under normal cell culture conditions. In 24 hours HUVEC cells formed a single spheroid (aggregate) within each well.

4.2.4 Sprouting assay

HUVEC spheroid aggregates were embedded in fibrin gels. Fibrin gels were prepared using fibrinogen (Sigma), aprotinin (Sigma), human Factor XIII (gift from Dr. Balazs Dome, National Koranyi Institute for TB and Pulmonology, Budapest, Hungary) and thrombin (Sigma) following the protocol of Helm et al. [136, 137]. Briefly, fibrinogen solution (3 mg/ml) containing aprotinin and Factor XIII was made and transformed into fibrin gel by supplementing thrombin enzyme. Before gelation, HUVEC aggregates were added to the fibrinogen solution and the mixture was transferred into 6 mm diameter circular wells. Three of such polylactic acid (PLA) wells were fused filament deposition- (3D-) printed into tissue culture dishes (Greiner) using a suitably modified Ultimaker Original printer. After gelation, fibrin gels filling the circular wells were covered with assay medium and kept under normal cell culture conditions within a microscope stage-top incubator.

4.2.5 Time lapse imaging

Time-lapse recordings of the fibrin gel-embedded HUVEC aggregates were performed on a Zeiss Axio Observer Z1 inverted microscope with 10x Plan Neofluar objective. The microscope was equipped with a Zeiss Axiocam MRM CCD camera and a Marzhauser SCAN-IM powered stage. Cultures within tissue culture Petri dishes (Greiner) were kept in a stage-mounted incubator providing 37°C and a humidified 5% CO₂ atmosphere. Stage positioning, focusing and image collection were controlled by Zeiss Axiovision 4.8 software and a custom experiment manager software module. Phase contrast images were collected every hour from each microscopic field for durations up to 72 hours.
4.2.6 Image analysis

Images recorded by time-lapse microscopy were analyzed by a custom-made segmentation algorithm using the NumPy [138], Scipy [139–141], Scikit-Image [142] and OpenCV [143] modules of Python 2.7. To identify the initial spheroid aggregates, the first frame of each image sequence was segmented by one of the following four methods: (i) Otsu’s thresholding, (ii) manual thresholding, (iii) Hough Circle Transform, (iv) adaptive thresholding. After segmentation, aggregates were reconstructed by using basic morphological operations and removal of small connected components. To detect sprout formation, images were preprocessed by extracting the background using a Gaussian mixture-based background segmentation algorithm [144, 145]. The resulting foreground image was blended with the reconstructed image of the initial aggregate, then basic morphological operations and small object removal were applied to reduce noise and to remove sprouts belonging to other aggregates. The sprouting spheroid was identified as the largest cluster of connected pixels. The image processing tools are available at https://github.com/doraelakatos/sprout-density-analyzer.git.

4.2.7 Sprout density

To quantify sprout density around the spheroids, we calculated the radial density profile $\rho_0(r,t)$ for each frame as:

$$\rho_0(r,t) = \frac{A_{\text{sprout}}(r + \delta r, r - \delta r, t)}{A_{\text{ring}}(r + \delta r, r - \delta r)},$$ (4.1)

where $r + \delta r$ and $r - \delta r$ are the outer and inner radii of a ring, respectively. The area of the ring is $A_{\text{ring}} = 4\pi r \delta r$. Within this ring, at a certain time $t$ the area occupied by sprouts is denoted by $A_{\text{sprout}}(r + \delta r, r - \delta r, t)$. The ring has the same center as the minimal enclosing circle of the initial spheroid aggregate, and we choose $\delta r = 5 \mu m$. To eliminate differences due to the initial conditions, we normalized the radial density profiles as

$$\rho(r,t) = \rho_0(r,t) - \rho_0(r,t_0),$$ (4.2)

where $\rho_0(r,t_0)$ is the density profile of the first frame.
4.3 Models

4.3.1 Reaction-diffusion

For simplicity, we consider only a single VEGF isoform, VEGF-A\(_{165}\), that readily binds the extracellular matrix (ECM) [20, 112, 146]. We investigate a scenario in which a vascular sprout invades an area where VEGF is immobilized by the ECM, its concentration is spatially uniform (\(V_0\)), and its availability as an extracellular signaling ligand is limited only by binding cell-secreted sVEGFR1 molecules. In our model VEGF can be either active (\(V\)) or inactivated by forming a complex with sVEGFR1 (\(R_b\)), thus

\[ V = V_0 - R_b \] (4.3)

holds. Active VEGF concentration (\(V\)) is determined by the local concentrations of free (\(R_f\)) and bound (\(R_b\)) sVEGFR1 according to the following reaction-diffusion dynamics:

\[
\begin{align*}
\partial_t V &= -\partial_t R_b = -k_{on}VR_f + k_{off}R_b + \gamma R_b, \\
\partial_t R_f &= -k_{on}VR_f + k_{off}R_b - \gamma R_f + \Gamma + D_R \nabla^2 R_f,
\end{align*}
\] (4.4, 4.5)

where \(k_{on}\) and \(k_{off}\) are the association and dissociation rates of the VEGF-sVEGFR1 complex, and \(D_R\), \(\gamma\) and \(\Gamma\) denote the diffusivity, degradation and the local secretion rate of sVEGFR1, respectively, and \(\partial_t\) represents the partial derivative with respect to time. For simplicity we assume that the degradation rate of sVEGFR1 is the same irrespective of forming a complex with VEGF, and its secretion rate is uniform \(\Gamma\), in areas occupied by cells and zero elsewhere (Fig. 4.1).

As the kinetics of receptor-ligand binding and complex dissociation is much faster than changes in the total amount of the protein, we can apply quasi steady state approximation (QSSA). Since complex formation equilibrates rapidly, it can be assumed as a steady-state process:

\[ k_{on}R_f V = k_{off}R_b, \] (4.6)

and the association and dissociation rates, \(k_{on}\) and \(k_{off}\) respectively, can be replaced by the equilibrium constant:

\[ K = \frac{k_{on}}{k_{off}}. \] (4.7)
Fig. 4.1. Model of sVEGFR1 driven vascular pattern formation. The concentration of VEGF (blue), immobilized by the ECM, is considered to be spatially uniform in the vicinity of the endothelial cell-covered area (yellow). The motility and proliferation of endothelial cells are promoted by the locally available VEGF via their cell surface receptors, VEGFR2 (green). Endothelial cells secrete a diffusive repressor, sVEGFR1 (red), that binds and inactivates VEGF. Thus, the concentration of active VEGF forms a gradient pointing away from endothelial cells (yellow arrow). As a protruding tip senses higher concentration and steeper gradients of active VEGF, it expands more rapidly, and further enhances its extension.

Thus, the concentration of bound sVEGFR1 can be expressed as a function of the concentration of free sVEGFR1 as

$$R_b(R_f) = \frac{KR_f V_0}{1 + KR_f}, \quad (4.8)$$

and the time derivative can be written as

$$\partial_t R_b = \frac{KV_0}{(1 + KR_f)^2} \partial_t R_f. \quad (4.9)$$

By introducing the total receptor concentration ($R$) as

$$R = R_f + R_b, \quad (4.10)$$
Chapter 4. Role of soluble VEGFR1 signaling in vascular patterning

and substituting it into the subtraction of Eq. (4.4) from Eq. (4.5) yields the time derivative of the total receptor concentration:

$$\partial_t R = D_R \nabla^2 R_f + \Gamma - \gamma R. \tag{4.11}$$

Replacing Eqs. (4.8-4.9) into this equation gives the QSSA dynamics of the freely diffusible receptor:

$$\partial_t R_f = \left[1 + \frac{KV_0}{(1 + KR_f)^2}\right]^{-1} \left[D_R \nabla^2 R_f + \Gamma - \gamma R_f \left(1 + \frac{KV_0}{1 + KR_f}\right)\right]. \tag{4.12}$$

We couple the dynamics of sVEGFR1, Eq. (4.12), to vascular patterning by the following assumptions. First, the secretion rate of sVEGFR1 is uniform for each cell. Second, the extension of the cell-covered area is promoted – by a combination of chemotaxis, increased motility and proliferation of endothelial cells – by the locally available VEGF. According to Eqs. (4.3) and (4.8), the VEGF concentration can be determined as the function of unbound inhibitor:

$$V = \frac{V_0}{1 + KR_f}. \tag{4.13}$$

By substituting the total receptor concentration into Eq. (4.8), we obtain a quadratic equation for \(R_b\) as a function of \(R\) with two positive real roots:

$$R_b = \frac{1 + KR + KV_0 \pm \sqrt{(1 + KV_0 + KR)^2 - 4K^2RV_0}}{2K}. \tag{4.14}$$

Since VEGF can be either active or inactivated by sVEGFR1 and its concentration is spatially uniform, substitution of Eq. (4.3) yields only one positive solution for \(V\):

$$V = \frac{KV_0 - 1 - KR + \sqrt{(1 + KV_0 + KR)^2 - 4K^2RV_0}}{2K}. \tag{4.15}$$

The binding affinity, \(1/K\), between VEGF and the decoy receptor sVEGFR1 is in the range of 2-30 pM [127, 147–149]. Accordingly, we set our concentration unit as \(c_0 = 20\) pM, and explored \(1/K\) values between 6 – 20 pM. The steady-state sVEGFR1 concentration in culture media conditioned by a monolayer of HUVEC cells was reported in the range of 100 – 200 pM, i.e., 5-10 \(c_0\) [150, 151]. Our choice of simulation time unit, which must be much smaller than the time scale of patterning, is \(t_0 = 10\) s. We set the parameter \(\gamma\) characterizing spontaneous protein degradation in the range
4.3. Models

4 \cdot 10^{-5} / t_0 \text{ to } 10^{-3} / t_0, \text{ corresponding to a lifetime range between 3 h and 3 days. The lifetime of most proteins falls in this range [152] and is also consistent with data characterizing the accumulation of sVEGFR1 in HUVEC culture [150, 151]. In a spatially homogeneous steady-state environment (}\Gamma = \Gamma_*\text{) lacking binding partners or cellular uptake, Eq. (4.5) yields } \gamma = \Gamma_* / R_f. \text{ We selected the production rate } \Gamma_* \text{ in the range of } 10^{-3} - 10^{-2} c_0 / t_0, \text{ corresponding to 7 - 70 pM/h. The experimental estimate for HUVEC monolayers is } \Gamma_* = 5 - 15 \text{ pM/h [150, 151]. Our choice of } \Gamma_* \text{ and } \gamma \text{ parameters thus yields steady-state inhibitor concentrations in the range of } 3 - 7 c_0, \text{ comparable with experimental data (5-10 } c_0 [150, 151]).}

The diffusion parameter of sVEGFR1 is not known in dense gels or tissues. As the molecule exhibits heparin binding sites [153–156], we expect that it can bind to the extracellular matrix environment, and thus assumed rather low diffusivity. We used } D_R \text{ in the range of } 10^{-8} - 10^{-7} \text{ mm}^2/\text{s, the value reported for bFGF diffusion in basement membrane ECM [44] and three-four orders of magnitude less than diffusion coefficients of proteins in aqueous solutions. These values reflect that the diffusion of sVEGFR1 involves repeated binding and unbinding events which substantially restrict the molecule’s free diffusion [29]. The natural length scale of our simulation is that of the cells, thus we choose } \ell_0 = 1 \mu\text{m. This choice leads to } D_R \approx 0.1 - 1 \ell_0^2 / t_0, \text{ when expressed using the natural units of our simulations. The commonly used minimal concentration of exogenous recombinant VEGF to elucidate motility or proliferative effects on endothelial cells is } 5 \text{ ng/ml} = 100 \text{ pM} = 5 c_0 \text{ (R&D Systems, Catalog # 293-VE/CF). ECM-bound VEGF, however is much more potent [125, 157], exerting signaling activity at concentrations as low as } 0.1 \text{ ng/ml} = 2.5 \text{ pM} = 0.1 c_0. \text{ The initial, spatially uniform, VEGF concentration was therefore chosen in our simulations as } V_0 = c_0 = 20 \text{ pM.}

4.3.2 Lattice model

To express the interdependence of vascular growth and sVEGFR1 diffusion, we augmented a simple lattice model with the reaction-diffusion equations (4.12) and (4.13). In this model the state } \sigma \text{ of each lattice site } x \text{ can be either empty } (\sigma(x) = 0) \text{ or occupied by cells } (\sigma(x) = 1). \text{ The inhibitor, sVEGFR1, is produced with a rate } \Gamma_* \text{ at lattice sites occupied by cells:}

\[ \Gamma(x) = \Gamma_* \sigma(x). \]  

(4.16)

In the model, cells spread by the following mechanism. In each elementary step of the stochastic simulation, two neighbor sites } x \text{ and } x' \text{ are selected randomly. If } x \text{ is occupied by cells, they can spread to } x' \text{ (by leaving } x \text{ also occupied) with a probability...}
p. The spreading process represents both VEGF-induced proliferation and motility. During a sufficiently short time period $dt$, the spreading probability is a linear function of $\chi(V)$, the VEGF sensitivity function as

$$p(x \rightarrow x') = \nu \chi(V(x')) dt$$  \hspace{1cm} (4.17)

where

$$\chi(V) = \begin{cases} V - V_* & \text{for } V > V_* \\ 0 & \text{otherwise.} \end{cases}$$  \hspace{1cm} (4.18)

Parameter $\nu$ characterizes the strength of VEGF response, both proliferative and motogenic.

Eq. (4.18) assumes a simple linear behavior above a threshold $V_*$. The existence of a threshold is a crude approximation of the nonlinear behavior of binding between VEGF and its main cell surface receptor, VEGFR2, its neuropilin co-receptors, and subsequent dimerization of the receptors to initiate intracellular signal transduction reactions [122, 158]. In our simulations we used the threshold $V_* = c_0/2$, approximately corresponding to a 10 pM affinity between VEGF and its cell surface receptor complex. Our choice of $\nu = 0.025/c_0$ corresponds to a typical spreading speed of $\nu c_0 \ell_0 / t_0 \approx 10 \mu m/h$ – similar to the typical speed of endothelial cell motility observed in culture [159–161].

### 4.3.3 Phase-field model

In the phase-field formulation the discrete state variable $\sigma$ is replaced by a density-like variable, $\phi(x, t)$. Cell-free and cell-occupied areas are distinguished by the values $\phi = 0$ and $\phi = 1$, respectively. Instead of a sharp boundary separating these values, however, phase-field models operate with transition zones where the phase value changes gradually between the two extremes.

The simplest, commonly used [162–164] equation determining the time evolution of the phase-field $\phi$ is

$$\partial_t \phi = -\delta_\phi F = D_\phi \nabla^2 \phi - \partial_\phi f,$$  \hspace{1cm} (4.19)

where $F$ is a functional, analogous to the free energy of a physical system, written in the form of

$$F(\phi) = \int_\Omega \left[ \frac{D_\phi}{2} |\nabla \phi|^2 + f(\phi, ...) \right] dx.$$  \hspace{1cm} (4.20)

Integration in Eq. (4.20) encompasses the entire system and the free energy density function $f(\phi, ...)$ encodes the relative stability of the two phases for various values
4.3. Models

of external parameters [165–168]. The parameter $D_\phi$ needs to provide a microscopic dynamics compatible with the time and length scales of the patterning process. Thus, gradients in $\phi$ need to be orders of magnitude steeper than those characteristic for the external field, hence we have chosen $D_\phi = 5 \cdot 10^{-3} D_R$.

To describe VEGF-induced vascular growth, we cast $f$ in the form of

$$f(\phi, V) = \frac{1}{\tau}(g(\phi) - \mu \chi(V)p(\phi)), \quad (4.21)$$

where $\tau$ sets the characteristic time scale of changes within the phase-field and parameter $\mu$ is analogous to $\nu$, the parameter characterizing VEGF sensitivity in the lattice model.

The double-well potential

$$g(\phi) = \frac{1}{4}\phi^2(1 - \phi)^2 \quad (4.22)$$

exhibits two minima at $\phi = 0$ and $\phi = 1$. The expression $\mu \chi(V)p(\phi)$ modulates the minima of $g$ and thus characterize – in the presence of VEGF – the preference of the cell-covered state over the cell-free state. The interpolating function $p(\phi)$ satisfies $p(0) = 0$ and $p(1) = 1$, and its usual choice [165, 167] is

$$p(\phi) = \phi^3(10 - 15\phi + 6\phi^2). \quad (4.23)$$

Substitution of Eqs. (4.22 - 4.23) into the free energy density function Eq. (4.21) and derivation with respect to $\phi$ yields the phase-field equation for expanding endothelial cells:

$$\frac{\partial \phi}{\partial t} = D_\phi \nabla^2 \phi + \frac{1}{\tau}\phi(1 - \phi) \left[ \phi - \frac{1}{2} + 30\mu \chi(V)\phi(1 - \phi) \right]. \quad (4.24)$$

To set the parameters $\tau$ and $\mu$ of the phase-field model, we consider a simpler, one dimensional version of the problem – the steady-state propagation of a straight interface with a velocity $v$. In this scenario the phase-field satisfies the advection equation

$$v \partial_x \phi + \partial_t \phi = 0, \quad (4.25)$$

and replacing the time derivative of Eq. (4.19) with Eq. (4.25) yields

$$0 = D_\phi \nabla^2 \phi + v \partial_x \phi + \partial_\phi f(\phi, V). \quad (4.26)$$
Next we examine how the speed of propagation $v$ depends on a pre-determined uniform VEGF response parameter $\mu \chi(x, t) = \chi_0$. Close to the interface ($\phi(x) \approx \frac{1}{2}$) the derivative $\partial_\phi f$ is dominated by the VEGF response term of Eq. (4.24) as

$$\partial_\phi f (\phi) \approx \frac{30\chi_0}{\tau} \phi^2 (1 - \phi)^2.$$  \hspace{1cm} (4.27)

We rescale our variables by a positive factor $s$ as

$$\hat{x} = \frac{x}{s}$$  \hspace{1cm} (4.28)

and define the rescaled field $\hat{\phi}$ as

$$\hat{\phi} (\hat{x}) = \phi(x).$$  \hspace{1cm} (4.29)

Expressing Eq. (4.26) in terms of the rescaled variables and using the approximation (4.27), we obtain

$$0 = D_\phi \hat{\phi}'' \left( \frac{X}{s} \right) + s \hat{\phi}' \left( \frac{X}{s} \right) + \frac{30}{\tau} \hat{\chi}_0 s^2 \hat{\phi}^2 \left( \frac{X}{s} \right) \left( 1 - \hat{\phi} \left( \frac{X}{s} \right) \right)^2.$$  \hspace{1cm} (4.30)

Thus, if $\hat{\phi}(\hat{x})$ is a solution of Eq. (4.30) with velocity $\hat{v}_0$ and VEGF response $\hat{\chi}_0$, then $\phi(x) = \hat{\phi} \left( \frac{x}{s} \right)$ is also a solution with velocity $v = s \hat{v}$ and VEGF response $\chi_0 = s^2 \hat{\chi}_0$. Thus,

$$v = s \hat{v} = \frac{\hat{v}_0 \sqrt{\chi_0}}{\sqrt{\chi}_0},$$  \hspace{1cm} (4.31)

yielding

$$v \sim \sqrt{\chi_0}.$$  \hspace{1cm} (4.32)

To validate and calibrate the scaling relationship (4.32), we performed simulations of the phase-field model equation (4.24) with a fixed, spatially uniform VEGF response $\chi_0$. The initial condition corresponds to the setting considered in our analysis as

$$\phi(x, y) = \mathcal{H}(-x),$$  \hspace{1cm} (4.33)

where $\mathcal{H}()$ is the step function: $\mathcal{H}(x) = 1$ for $x > 0$ and zero otherwise. The boundary position, $X$, was defined as $\phi(X, 0) = 1/2$. In agreement with Eq. (4.32), the velocity $v$ of the advancing front for our choice of $D_\phi$ and $\tau$ could be well fitted as

$$v = a \sqrt{\chi_0},$$  \hspace{1cm} (4.34)
for a wide range of the prescribed driving force $0 \ 1/c_0 \leq \chi_0 \leq 16 \ 1/c_0$ with $a = (7.6 \cdot 10^{-3} \pm 1\%) \ \ell_0/t_0 \approx 5.5 \mu m/h$. To obtain spreading speeds similar to that of the lattice model (10 $\mu m/h$) at typical VEGF concentrations $V - V_\ast = c_0$, we set $\mu = 13.3 \ 1/c_0$. Substitution of Eq. (4.18) into Eq. (4.34) then yields the propagation speed of the boundary as a function of $V_X$, the concentration of VEGF at the boundary:

$$v = \begin{cases} 
  a\mu \sqrt{V_X - V_\ast} & \text{for } V_X > V_\ast, \\
  0 & \text{otherwise.}
\end{cases}$$

(4.35)

4.4 Results

4.4.1 Endothelial sprouting modulated by VEGFR1

In order to obtain time-resolved quantitative data about the role of sVEGFR1 in vascular sprouting, we performed experiments with human umbilical cord vein endothelial cells (HUVECs) and commercially available function blocking antibodies that bind both the soluble and membrane-bound form of VEGFR1. Our sprouting assay consisted of HUVEC aggregates embedded in fibrin gel and supplemented with exogenous VEGF. During a time course of a day, endothelial cells left the aggregates in multicellular sprouts and created a vascular structure (Fig. 4.2A, A). Physical cross sections revealed lumen-forming vessels with a diameter comparable to the size of an individual endothelial cell, within the range of $10 - 30 \mu m$. While most vascular segments grew straight during our one day long time-lapse observation period, some formed branched structures. After six days in culture, however, a branching network is more prominent (Fig. 4.3).

Vascular sprouts grown in the presence of function blocking VEGFR1 antibodies were denser (exhibited more sprouts per unit volume) than those grown in control cultures (Fig. 4.2B). To quantitatively characterize vascular sprouting in this experimental model system, we segmented the area covered by cells by an image processing algorithm. The segmented image sequence was evaluated in terms of normalized radial density profiles $\rho(r, t)$: for a sequence of concentric rings with various radii $r$, we determined the fraction of cell covered area within each ring (Fig. 4.2C-D). To focus on sprouts, density profiles were normalized by removing the contribution of the original aggregate. The time-dependent radial density profiles reveal that irrespective of the presence of antibodies, the sprouts expanded with a steady rate of 70 $\mu m/day = 3 \mu m/h$ (Fig. 4.4B), which is substantially slower than the speed of individual endothelial cells in culture (10-20 $\mu m/h$, [159–161]). The density of vascular
Fig. 4.2. Sprouting assay using HUVEC spheroids in fibrin gel. Representative phase-contrast images from time-lapse recordings (A) of untreated control cultures (A), and of cultures exposed to function blocking antibodies against VEGFR1 at a concentration of 20 µg/ml (B). Exposure to the antibody reduces the interbranch distance and yields denser sprouts. Vascular components recognized by the image analysis algorithm are colored blue. Representative radial density profiles, $\rho(r, t)$, are shown as a function of radius, for untreated (C) and a-VEGFR1 treated (D) sprout structures. The color saturation of the lines indicates the in vitro age of the culture between 10 and 30 hours, plotted for each frame.

branches, however, is higher in the presence of function blocking VEGFR1 antibodies. To quantitatively characterize this phenomena, we determined $\rho_{\text{max}}(t)$, the maximal value of $\rho(r, t)$, for each time-lapse frame $t$. For untreated cultures, vascular sprouts cover 40-50% of the area close to the aggregate. In contrast, the presence of VEGFR1 antibodies increases the maximal coverage to 60%. The density difference between treated and untreated cultures is visualized in Fig. 4.4A, by subtracting the population average $\langle \rho_{\text{max}}^{\text{control}}(t) \rangle$ of control cultures from the population average $\langle \rho_{\text{max}}^{\text{treated}}(t) \rangle$ of treated cultures for each time point $t$. Data pooled from 19 distinct sprout networks
4.4. Results

Fig. 4.3. Sprouting assay using HUVEC spheroids in fibrin gel. After six days in culture a branching vascular structure develops.

supports the observation that the sprout network is denser in the presence of function blocking VEGFR1 antibodies throughout the entire recorded time period. The cumulative distribution function of normalized sprout densities within each sprout systems is shown in Fig. 4.5A. Statistical significance of the differences were established by t-tests ($p < 0.05$).

4.4.2 Simulations

To explore the reaction-diffusion guided patterning mechanism formulated in our models, we performed computer simulations in two dimensions. The reaction-diffusion equation Eq. (4.12) was coupled to either the phase-field equation (Eq. (4.24)) or to the spreading probability (Eq. (4.17)) in the lattice model. Discretized equations were solved on a uniform, $N \times N$ grid by employing the forward Euler method. The lattice size was $\ell_0$ and $\ell_0/2$ in the lattice and the phase-field simulations, respectively. To avoid numerical instabilities, the time step $\Delta t = 0.1\ell_0$ was chosen in accordance with stability conditions. In case of the lattice model, the continuous variables were updated after $N^2$ elementary steps. To suppress the inherent anisotropy of the square lattice in phase-field simulations, we introduced a multiplicative “quenched” noise $0 < n(x) < 2$ with unit mean and $1/\sqrt{3}$ standard deviation in the chemotactic sensitivity as $\mu(x) = n(x)\mu$ [169].

As experimental data leave a substantial uncertainty about the parameter values, here we present results obtained with distinct choices (see Table 4.1) for the lattice and
Chapter 4. Role of soluble VEGFR1 signaling in vascular patterning

Fig. 4.4. Soluble VEGFR1 increases sprout density both in experiments and in simulations. A): Density difference $\langle \rho_{\text{treated}}(t) \rangle - \langle \rho_{\text{control}}(t) \rangle$ values demonstrate that cultures where VEGFR1 was blocked develop a denser sprout network than untreated cultures do. The lines show the average difference in density as a function of the age of culture $t$ – obtained from two distinct experiments containing $n = 8$ and $n = 11$ HUVEC spheroids. B): Sprout expansion in experiments is insensitive to the presence of VEGFR1 antibodies. Sprout extent was established from normalized radial density profiles ($\rho(r, t)$, see Fig. 4.2) as the radius $R$ where $\rho(R, t)$ is the half of the maximal density $\rho_{\text{max}}(t)$. Blue and red lines represent control and a-VEGFR1 treated cultures, respectively from two distinct experiments. Irrespective of the presence of antibodies, the sprouts expanded at similar rates, in the range of $2.8 – 3.4 \, \mu\text{m}/\text{h}$. C): Density difference in lattice simulations where soluble (red), membrane-bound (green) or both (yellow) population of VEGFR1 was blocked. In the simulations a 10% increase in the degradation rate (Eq. 4.36) represents the presence of a-sVEGFR1 antibodies, and a 10% shift in the VEGF response curve (Eq. 4.37) represents the inhibition of cell surface receptors. Each line indicates mean values obtained from four distinct simulations. D): Sprout expansion in lattice simulations was calculated by the same method used to evaluate experimental results. Expansion speed baseline (blue): $1.7 \pm 0.01 \, \mu\text{m}/\text{h}$, sVEGFR1-blocked (red): $1.9 \pm 0.01 \, \mu\text{m}/\text{h}$, mVEGFR1-blocked (green): $2.4 \pm 0.02 \, \mu\text{m}/\text{h}$, both population blocked (yellow): $2.49 \pm 0.02 \, \mu\text{m}/\text{h}$. Shaded areas represent the standard error of the mean, calculated from individual spheroids treated in identical manner (A, B) and distinct simulations (C, D).
4.4. Results

Fig. 4.5. Cumulative distribution of radial density profiles. A): Normalized sprout densities \( \rho_{\text{max}}(t) / \langle \rho_{\text{max,control}}(t) \rangle \), obtained for each spheroid at in vitro age \( t = 20 \) h. Sprout densities were normalized by the average of control sprout meshwork densities in each experiment. The differences are significant based on a t-test \( (p < 0.05) \). B): Normalized sprout densities, each obtained from \( n = 5 \) independent simulations of the lattice model. Blue and red lines represent control and a-VEGFR1 treated cultures, respectively.

phase-field simulations. Despite using different parameter sets and different mathematical model formulations, the reaction-diffusion mechanism can produce branching patterns in a robust manner (Figs. 4.6, 4.7). Moreover, using the same set of parameters yields similar patterns in both implementations (Fig. 4.8).

Thus, model simulations indicate the presence of tip-splitting instability as the initial cluster of cells expands into a branching pattern (Fig. 4.9). The secretion of sVEGFR1 inactivates the VEGF in the vicinity of the sprouts, and the threshold condition Eq. (4.18) is required to freeze those boundaries which are behind the expanding envelope of the branch tips. When the threshold concentration \( V_\ast \) is not substantial enough (i.e., \( V_\ast \ll c_0 \)), vascular segments continue to widen even far behind the branch tips and thus create a compact structure.

The profile of the self-generated VEGF gradient was determined – along the longitudinal direction – at selected branch tips (Fig. 4.10). We also determined the profiles corresponding to the Mullins-Sekerka system, in which an external diffusive field (i.e., VEGF) drives the propagation of the boundary [170] (Fig. 4.11). The obtained profiles indicate that the gradient is 2-3 times steeper in the reaction-diffusion system than in the Mullins-Sekerka system with comparable parameters.

The typical branch width of the simulations is in the range of 10-30 \( \mu \)m, in accord with the experimentally observed values. Branches expand with a speed of \( 4.8 \cdot 10^{-2} - 9.6 \cdot 10^{-2} \ell_0 / t_0 \approx 1.7 - 3.4 \mu \text{m/h} \), values close to the experimental observations (3 \( \mu \text{m/h} \)). The interbranch distance as well as the diffusion length of sVEGFR1
are both $\sim 80 \ell_0$, consistent with the theoretical estimate $\sqrt{D_R/\gamma}$. Indeed, both increasing $D_R$ and decreasing $\gamma$ leads to sparser branches (Fig. 4.12). The typical magnitude of VEGF gradients that develop in the model is $7 - 12 \cdot 10^{-3} c_0/\ell_0 \approx 140 - 240$ pM/mm = 5-10 ng/(ml mm) – the same magnitude that the endothelial cells are able to detect in cell culture experiments [171].

Our parameter choice yields an average cell coverage of 20% in the lattice model – which can be increased by faster sVEGFR1 degradation, but is unaffected by changing its diffusivity $D_R$. In general, our simulations suggest that when parameters favor expansion (decreased threshold concentration, increased chemotactic response, etc) the branches are wider, and eventually they merge into a smooth propagating front. When parameters are less conductive to cell expansion, branches became thinner, and below a threshold, expansion ceases (Figs. 4.13 and 4.14).

### 4.4.3 Model validation

We validated our model by comparing it to experimental sprouting data obtained in the presence and absence of VEGFR1-blocking antibodies. While the antibody is thought to block both the membrane-bound and diffusive form of the inhibitor, the computational model can predict the consequences of blocking either form of the receptor.

To represent in our models the experimental perturbation of function blocking antibodies interacting with the soluble receptor, we altered the VEGFR1 degradation rate parameter $\gamma$ as VEGFR1 can be inactivated both by degradation and by dimerization with a high affinity antibody. Since in the experiments antibodies were given in large excess, its concentration is approximately constant, $A$. Thus, inactivation of sVEGFR1 is given by

$$\partial_t R = - \left( \gamma + k_{on}^A A \right) R = -\gamma' R,$$

where $k_{on}^A$ is the association rate of the sVEGFR1-antibody complex and $\gamma'$ is the increased degradation rate ($\gamma' > \gamma$). In agreement with experimental observations, when $\gamma$ is increased both the lattice and the phase-field model predicts denser sprouts (Fig. 4.15) and a slight (less than 10% increase) in expansion speed (Fig. 4.4).
4.4. Results

Fig. 4.6. Time development of the lattice model. The area occupied by the cells and the concentration fields of the free growth factor (VEGF) and of the inhibitor (sVEGFR1) are shown in the left, middle and right columns, respectively. The state of the simulation is shown at the beginning (top) and after 20000, 40000 and 60000 (bottom) time units, corresponding approximately to 2, 4.5 and 7 days in culture. Parameters of the $800 \times 800$ lattice simulation are summarized in Table 4.1.
Fig. 4.7. Time development of the phase-field model. As in Fig. 4.6, the area occupied by the cells and the concentration fields of the free growth factor (VEGF) and of the inhibitor (sVEGFR1) are shown in the left, middle and right columns, respectively. The state of the simulation is shown at the beginning (top) and after 5000, 12500 and 20000 (bottom) time units, corresponding approximately to 14, 35 and 55 hours of development. Parameters of the 1200 × 1200 lattice simulation with Δx = ℓ₀/2 mesh spacing are summarized in Table 4.1.
Fig. 4.8. Time development of the lattice model using the same set of parameters as in the phase-field model (Fig. 4.7). The area occupied by the cells, the concentration fields of the free activator (VEGF) and of the inhibitor (sVEGFR1) are shown in the left, middle and right columns, respectively. The state of the simulation is shown at the beginning (top) and after 2000, 4000 and 6000 (bottom) time units, corresponding approximately to 1, 2 and 3 days in culture.
Chapter 4. Role of soluble VEGFR1 signaling in vascular patterning

Fig. 4.9. Tip-splitting event in the phase-field model. A): The area occupied by the cells are shown at every 1000\(t_0\) simulation steps (\(\sim 2.7\) h). B): The concentration field of the free activator (VEGF) is shown at three different time points. The thicker and the thinner black lines correspond to a VEGF contour (\(V = 0.65c_0\)) and to the boundary of the vasculature (\(\phi = 0.65\)) respectively. Red arrows point to sites of future branches: border segments that experience especially high VEGF concentrations.

Fig. 4.10. VEGF concentration profiles along a growing branch tip in phase-field models. Red and blue curves show VEGF concentrations in simulations where the activator or the inhibitor diffuses, respectively. Grey lines depict the phase field.
Fig. 4.11. Time development of the diffusing VEGF-driven phase-field model. Reaction-diffusion dynamics (Eqs. 4.4-4.5) has been modified that instead of a diffusing inhibitor (sVEGFR1) an external diffusive field (VEGF) drives the propagation ($D_R = 0$, $D_V = 1 \ell_0/t_0$). As in Fig. 4.7, the area occupied by the cells and the concentration fields of the free activator (VEGF) and of the inhibitor (sVEGFR1) are shown in the left, middle and right columns, respectively. The state of the simulation is shown at the beginning (top) and after 5000, 10000 and 15000 (bottom) time units, corresponding approximately to 14, 28 and 42 hours of development. Parameters of the $800 \times 800$ lattice simulation with $\Delta x = \ell_0/2$ mesh spacing are summarized in Table 4.1, except $\Gamma_* = 0.0012 c_0/t_0$. 
Fig. 4.12. The lattice model yields a dense branching morphology. The cumulative radial density (A, B) reaches a steady value, indicating a space-filling structure with a fractal dimension of 2. Morphologies are shown for three values of the diffusivity (C) and the degradation rate (D) of the inhibitor. The radial density profiles in panels (A, B) can be matched by the color code to the morphologies in panels (C, D). $\gamma_0 = 10^{-3} \, 1/t_0$.

Fig. 4.13. Parameter sensitivity of branching morphology in the lattice model. Panels show the area occupied by the cells after $60000 \, t_0$ in lattice simulation. In each simulation, one of the parameters summarized in Table 4.1 was altered, as indicated above each panel.
Fig. 4.14. Parameter sensitivity of branching morphology in the phase-field model. Panels show the area occupied by the cells after 14000 $t_0$ in phase-field simulation. In each simulation, one of the parameters summarized in Table 4.1 was altered, as indicated above each panel.

Fig. 4.15. Representative sprout density profiles in the computational models. Radial density profiles ($\rho(r,t)$) are shown for two representative lattice model (A, B) and phase-field (C, D) simulations. Like in Fig. 4.2, color saturation indicates time elapsed since the onset of simulation. According to Eq. (4.36), in the simulations an increased degradation rate ($\gamma'$) represents the presence of a-VEGFR1 antibodies. Blue and red lines correspond to simulations representing control cultures (i.e., parameters chosen according to Table 4.1) and a-VEGFR1 treated cultures, respectively.
To represent antibody binding to cell surface receptors, we shift the VEGF-response curve (4.18) by a factor \( \alpha \) as

\[
\chi(V) = \begin{cases} 
  \alpha V - V_* & \text{for } \alpha V > V_* \\
  0 & \text{otherwise.}
\end{cases}
\] (4.37)

Thus, we assume that competitive binding at the cell surface changes the EC-50 value of the VEGF dose-response curve. As the antibodies remove an inhibitory effect, we expect \( \alpha > 1 \). Again, simulations predict an increase in vascular density, with a larger increase in the speed of pattern propagation (Fig. 4.4).

The experimental findings and simulation results are thus in reasonable agreement. As in the experiments the increase in sprout growth speed is smaller than the increase in vascular density, we suggest that the effect of the antibodies are mainly exerted through their association with the diffusive form of VEGFR1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Experimental value</th>
<th>Reference</th>
<th>value in lattice model</th>
<th>value in phase-field model</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVEGFR1 parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>diffusivity (( D_{\phi} ))</td>
<td>( 10^{-7} ) mm(^2)/s</td>
<td></td>
<td>( 1 \ell_0^2/t_0 )</td>
<td>( 0.25 \ell_0^2/t_0 ) = ( 2.5 \times 10^{-8} ) mm(^2)/s</td>
</tr>
<tr>
<td>binding affinity to VEGF (1/K)</td>
<td>( 2 - 30 ) pM</td>
<td>[127, 147-149]</td>
<td>( 1/c_0 )</td>
<td>( 0.333 \ell_0^2/t_0 = 6.67 ) pM</td>
</tr>
<tr>
<td>secretion rate (( f_\text{sV} ))</td>
<td>( 5 - 15 ) pM/h</td>
<td>[150, 151]</td>
<td>( 10^{-2} c_0/t_0 ) = ( 72 ) pM/h</td>
<td>( 10^{-2} c_0/t_0 ) = ( 7.2 ) pM/h</td>
</tr>
<tr>
<td>degradation rate (( \gamma ))</td>
<td>( 0.03 - 0.8 ) 1/h</td>
<td>[152]</td>
<td>( 10^{-2} /t_0 ) = 0.36 1/h</td>
<td>( 4 \times 10^{-3} /t_0 = 0.01 ) 1/h</td>
</tr>
<tr>
<td>steady state concentration</td>
<td>100 - 200 pM</td>
<td>[150, 151]</td>
<td>( 3.5 c_0 ) = 70 pM</td>
<td>( 7 c_0 ) = 140 pM</td>
</tr>
<tr>
<td>VEGF parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cellular sensitivity (( \nu, \mu ))</td>
<td>( 10^{-2} \ell_0^2/t_0 )</td>
<td>[157, 172, 173]</td>
<td>( 0.5 c_0 ) = 10 pM</td>
<td>( 0.5 c_0 ) = 10 pM</td>
</tr>
<tr>
<td>threshold concentration (( V_* ))</td>
<td>( 2.5 ) pM</td>
<td>[157, 172, 173]</td>
<td>( 0.5 c_0 ) = 10 pM</td>
<td>( 0.5 c_0 ) = 10 pM</td>
</tr>
<tr>
<td>other simulation parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>phase-field diffusivity (( D_{\phi} ))</td>
<td>( 5 \times 10^{-3} \ell_0^2/t_0 )</td>
<td>[127, 174]</td>
<td>( 0.1 c_0 ) = 20 ( 1/t_0 )</td>
<td>( 0.1 c_0 ) = 20 ( 1/t_0 )</td>
</tr>
<tr>
<td>characteristic time scale (( \tau_0 ))</td>
<td>( 2 ) s</td>
<td></td>
<td>( 0.1 t_0 ) = 0.2 s</td>
<td>( 0.1 t_0 ) = 0.2 s</td>
</tr>
<tr>
<td>time step (( \Delta t ))</td>
<td>( 0.1 ) s</td>
<td></td>
<td>( 0.1 t_0 ) = 0.1 s</td>
<td>( 0.1 t_0 ) = 0.1 s</td>
</tr>
<tr>
<td>mesh spacing (( \Delta x ))</td>
<td>( 1 \ell_0 )</td>
<td></td>
<td>( 0.5 \ell_0 )</td>
<td>( 0.5 \ell_0 )</td>
</tr>
<tr>
<td>emergent properties</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>branch width</td>
<td>( 10.6 - 29.2 ) ( \mu ) m</td>
<td></td>
<td>( 5.8 - 13.7 \ell_0 ) = ( 5.8 - 13.7 ) ( \mu ) m</td>
<td>( 13.2 - 37.2 \ell_0 ) = ( 13.2 - 37.2 ) ( \mu ) m</td>
</tr>
<tr>
<td>branch growth speed</td>
<td>( 3.05 \pm 0.52 ) ( \mu ) m/h</td>
<td></td>
<td>( 4.8 \times 10^{-3} \pm 13 % \ell_0/t_0 = 1.73 \pm 0.22 ) ( \mu ) m/h</td>
<td>( 9.6 \times 10^{-2} \pm 5 % \ell_0/t_0 = 0.36 \pm 0.16 \mu ) m/h</td>
</tr>
<tr>
<td>VEGF gradient</td>
<td>( 50 ) pM/min &lt;</td>
<td>[127, 174]</td>
<td>( 2 \times 10^{-2} \times 10^{-3} \ell_0/t_0 = 140 \pm 14 ) pM/min</td>
<td>( 1.2 \times 10^{-2} \times 10^{-3} \ell_0/t_0 = 240 \pm 24 ) pM/min</td>
</tr>
<tr>
<td>VEGF gradient</td>
<td>( 7 \times 10^{-2} \times 10^{-3} \ell_0/t_0 = 140 \pm 14 ) pM/mm</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1. Parameters and emergent properties.

4.4.4 Analysis of front propagation

4.4.4.1 Concentration profile generated by a moving front

To better understand the simulation results, we consider a plane boundary propagating according to the locally available VEGF concentration. First, we determine the steady state concentration profile of the diffusive repressor field when the boundary propagates with a steady speed \( v \). In order to solve our reaction-diffusion system analytically, we simplify our problem and assume that (i) the inhibitor diffuses irrespectively whether it is in a complex or not, and (ii) the boundary is sharp. Under
these assumptions the total receptor concentration satisfies

$$\partial_t R = D_R \left( \partial_{xx} + \partial_{yy} \right) R - \gamma R + \Gamma. \quad (4.38)$$

Reflecting a sharp boundary located at $x = X_0(t)$, the secretion rate is given by

$$\Gamma(x, y) = \Gamma_* \mathcal{H}(X_0(t) - x). \quad (4.39)$$

To calculate the stationary planar front solutions of Eq. (4.38), translating with a steady velocity $v$, we introduce the co-moving coordinate $\zeta = x - vt$. The boundary position and the secretion rate then change to $X = X_0(t) - vt$ and $\Gamma(\zeta, y) = \Gamma_* \mathcal{H}(X(t) - \zeta)$, respectively. With appropriate choice of $t$, we can set the position of the planar front to $X(t) = 0$. The new spatial variable $\zeta$ transforms the partial time derivative as $\partial_t |_{x} = \partial_t |_{\zeta} - v \partial_\zeta |_{t}$. Hence, in this scenario the total receptor concentration $r_0$ and secretion rate $\Gamma_0(\zeta)$ can be written as

$$R(\zeta, y) = r_0(\zeta) \quad (4.40)$$

and

$$\Gamma(\zeta, y) = \Gamma_0(\zeta) = \Gamma_* \mathcal{H}(-\zeta), \quad (4.41)$$

respectively.

Since the front has a stationary profile in the co-moving frame, the time derivative vanishes and Eq. (4.38) reduces to an ordinary differential equation for $r_0$:

$$D_R \partial_{\zeta \zeta} r_0 + v \partial_\zeta r_0 - \gamma r_0 + \Gamma_0(\zeta) = 0. \quad (4.42)$$

It is convenient to construct $r_0$ from two distinct spatial components as

$$r_0(\zeta) = \begin{cases} r_0^{-}(\zeta) & \text{for } \zeta < 0 \\ r_0^{+}(\zeta) & \text{for } \zeta > 0 \end{cases}. \quad (4.43)$$

By solving the corresponding homogeneous differential equations, we obtain the following eigenvalues of the characteristic equation as:

$$\lambda_{+,-} = -v \pm \frac{\sqrt{v^2 + 4D_R\gamma}}{2D_R}. \quad (4.44)$$
Chapter 4. Role of soluble VEGFR1 signaling in vascular patterning

For a front propagating from left to right, the boundary and continuity conditions are

\[ r_0^- (\xi \to -\infty) = \frac{\Gamma_s}{\gamma}, \]
\[ r_0^+ (\xi \to \infty) = 0, \]
\[ r_0^- (\xi = 0) = r_0^+ (\xi = 0), \]
\[ \partial_\xi r_0^+ (\xi = 0) = \partial_\xi r_0^- (\xi = 0), \tag{4.45} \]

respectively. Conditions (4.45) allow to fully specify the solution for Eq. (4.42):

\[ r_0^- = \frac{\Gamma_s}{\gamma} \left( 1 - \frac{\lambda_-}{\lambda_- - \lambda_+} e^{\lambda_+ \xi} \right), \]
\[ r_0^+ = -\frac{\Gamma_s}{\gamma} \frac{\lambda_+}{\lambda_- - \lambda_+} e^{-\lambda_- \xi}. \tag{4.46} \]

The concentration of the inhibitor at the interface, \( r_X \), can be expressed as a function of the front velocity \( v \) as

\[ r_X (v) = r_0^- (\xi = 0) = \frac{\Gamma_s}{2\gamma} \left( 1 - \frac{v}{\sqrt{v^2 + 4D_R \gamma}} \right). \tag{4.47} \]

Substitution of \( r_X (v) \) into Eq. (4.15) yields an analytic estimate for the VEGF concentration at the interface as a function of \( v \) (Fig. 4.16) as

\[ V_{X,est} (v) = \frac{KV_0 - 1 - Kr_X (v) + \sqrt{1 + KV_0 + Kr_X (v)}^2 - 4K^2 r_X (v) V_0}{2K}. \tag{4.48} \]

To verify the correctness of the above picture, we also obtained \( V_X (v) \) from the full reaction-diffusion equations, i.e., without the simplifications of Eq. (4.38). In the co-moving frame the stationary profile of the free sVEGFR1 repressor satisfies

\[ 0 = D_R \partial_\xi R_f + v \left( 1 + \frac{KV_0}{(1 + KR_f)^2} \right) \partial_\xi R_f + \Gamma (\xi) - \gamma R_f \left( 1 + \frac{KV_0}{1 + KR_f} \right). \tag{4.49} \]

The boundary conditions for \( R_f \)

\[ R_f (\xi \to \infty) = 0, \tag{4.50} \]
\[ R_f (\xi \to -\infty) = -\frac{(KV_0 \gamma + \gamma - \Gamma K) + \sqrt{(KV_0 \gamma + \gamma - \Gamma K)^2 + 4\gamma K\Gamma}}{2\gamma K}. \tag{4.51} \]
Fig. 4.16. Velocity selection in the lattice (A) and phase-field (B) models. The red curves represent the expected propagation speed of the boundary for a prescribed value \( V_X \) of VEGF at the interface. The black and blue curves show the calculated (blue) and analytically estimated (black) \( V_X \) values for a pre-determined propagation speed of the boundary \( (v_0) \). Intersection of the red and blue curves determine the emergent propagation speed in the full model. In the absence of intersections, a planar front cannot propagate. By forming branches, the effective secretion rate of the inhibitor is decreased (dotted curves) allowing for the propagation of the leading tips of dense branching morphologies. \( V_X, est \) \( (v_0) \): analytic solution for \( V \) as the function of the propagation speed obtained from the differential equation of the simplified system; \( V_X, BVP \) \( (v_0) \): numerical solution of the reaction-diffusion equation of the free repressor in a co-moving frame with BVP solver; \( v_0(V_X) \): propagation speed of the boundary as the function of \( V \) obtained from simulation of the phase-field equation with uniform driving force; simulation: operating point obtained from the simulation of the full differential equation system.

specify a boundary value problem (BVP) for Eq. (4.49). The BVP was solved numerically for several \( v \) values and the free VEGF concentration at the boundary was obtained using Eq. (4.13). The resulting \( V_X, BVP \) \( (v) \) curves (Fig. 4.16) exhibit very similar behavior to our analytic estimate E. (4.48).

4.4.4.2 Velocity selection

A front propagating in a steady state needs to satisfy both Eq. (4.48) and Eq. (4.35), yielding a set of equations for two unknowns, \( v \) and \( V_X \). As Fig. 4.16 demonstrates, the \( v(V_X) \) and \( V_X(v) \) functions may or may not have intersections.

When intersection points exists, their \( v \) and \( V_X \) values specify possible steady states of moving fronts. However, when two intersection points are present, we expect the first one, i.e. the one corresponding to the smaller velocity, to be unstable. When VEGF concentration is slightly increased at the interface, this perturbation speeds up
the advancement of the boundary. In the vicinity of the first fixed point faster propagation further increases the VEGF concentration, thus the front will accelerate and reach the second, stable fixed point.

If there are no intersections of the $v(V_X)$ and $V_X(v)$ functions, propagating steady state fronts cannot arise in the system. Branching, however, allows a steady expansion of the vascular sprouts even under such conditions. As the inhibitor can diffuse away from the branches in the lateral directions, its concentration is reduced at the branch tips. The crude approximation of infinitely efficient azimuthal diffusive transport across the branches yields the previously analyzed planar front problem where the envelope of branch tips form the expanding boundary. The propagation of this coarse-grained sprout system is then subject to the same velocity selection conditions Eq. (4.48) and Eq. (4.35). However, as the inhibitor is produced only at a fraction $\rho$ of the surface, we need to scale its spatially uniform production factor as $\Gamma' = \Gamma / \rho$. Indeed, simulations performed using the full set of equations (4.24), (4.12) and (4.13) and started from an initial condition consistent with boundary conditions Eq. (4.51), confirm that our analysis correctly provides the selected velocity and VEGF concentration at the boundary (Fig. 4.16). Hence, when a compact front cannot propagate, branching can be seen as an effective way to reduce inhibitor production and restore the possibility of growth. The approximation of an infinitely efficient azimuthal transport cannot hold when the sprout structure is so sparse that the characteristic distance between the branches is larger than the diffusion length. At this point the inhibitor concentration at the tip of the branches cannot be lowered further by increased distance between the branches, hence for high enough inhibitor secretion branching structures will not be able to grow.

### 4.4.5 Generalizations of the model

The above results establish that a diffusing inhibitor (sVEGFR1) secreted by the cells can generate gradients of an activator (VEGF). The interplay between the diffusive fields and the moving source of the inhibitor can generate tip-splitting instabilities, reminescent of the multicellular sprouting behavior of endothelial cells in culture. Our simplified model, however, does not take into account several known aspects of endothelial biology. There is good evidence that the endothelial cells are not uniform: one can distinguish a population of leader cells with specific molecular and functional features[48, 175, 176] including the difference in sVEGFR1 secretion [131]. Endothelial cells are also known to internalize and eliminate VEGF [22, 177, 178], thereby altering the reaction-diffusion system considered in this work. To explore how these effects
may change the basic patterning mechanism, we augmented our model with two sets of rules.

4.4.5.1 Spatial variation in sVEGFR1 secretion

Without explicitly representing the tip cell selection process in our models, we introduced spatial inhomogeneity by restricting inhibitor secretion in areas recently occupied by the cells. Thus, we modify Eq. 4.16 to describe inhibitor secretion at time $t$ as

$$\Gamma(x) = \begin{cases} 
\Gamma_0 \sigma(x) & \text{for } t > t_0(x) + T_* \\
0 & \text{otherwise.}
\end{cases}$$

(4.52)

where $t_0$ denotes the time lattice site $x$ was occupied, and the $T_* \geq 0$ threshold is an adjustable parameter. For $T_* = 0$ we recover our original model. Increasing $T_*$ excludes secretion from increasingly larger areas beyond the leading edge of the sprouts.

Simulations with a small excluded area ($T_* = 300t_0$) does not alter the dynamics substantially (Fig. 4.17A-C). Simulations with $T_* = 500t_0$ yield branches and excluded areas $\approx 10 - 20 \mu m$ wide, corresponding to multicellular sprouts 1-2 cells wide headed by a single tip cell. In this scenario the branches are wider, but are more sparse, and the overall density of the cell-covered area is decreased (Fig. 4.18). Further increase in $T_*$ yields $\approx 100 \mu m$ wide branches, each led by an area corresponding to 10-100 leader cells. In this limit the expansion becomes oscillatory: the lack of inhibitor production results in a quick expansion of a compact cluster. As the inhibitor production turns on, the extended spatial production of the inhibitor yields a VEGFR1 concentration high enough that it can choke the further expansion of the cells.

4.4.5.2 VEGF internalization by the cells

Internalization of VEGF by endothelial cells [22, 177, 178] is expected to alter the spatial distribution of sVEGFR1 by freeing up molecules that could have bound in the absence of internalization. To gauge the effect of VEGF internalization on the patterning mechanism, we keep track of $V_T$, the local total amount of VEGF

$$V_T = V + R_b,$$

(4.53)

and expand Eqs (4.4) as

$$\partial_t V = -k_{on}VR_f + k_{off}R_b + \gamma_R R_b - \gamma_V V \sigma,$$

(4.54)

$$\partial_t R_b = k_{on}VR_f - k_{off}R_b - \gamma_R R_b,$$

(4.55)
Chapter 4. Role of soluble VEGFR1 signaling in vascular patterning

Fig. 4.17. Time development of the lattice model with heterogeneous sVEGFR1 secretion. The inhibitor is not produced at the branch tips: cells that were occupied less than $T_s$ time ago. Cell-occupied areas are shown for different values of $T_s$ (A-C: $T_s = 300t_0$, D-F: $T_s = 500t_0$, G-I: $T_s = 1000t_0$) at three snapshots of the simulations. Insets show branch tips at higher magnification (1:4.2). Orange colors indicate areas where the inhibitor is not produced, i.e., the presumed location of tip cells.

where $\gamma_V$ denotes the internalization rate of VEGF, thus its internalization is assumed to follow first order kinetics, where cells are present ($\sigma = 1$). We performed simulations with the lattice model, using the quasi steady state approximation.

As expected, for large enough values of $\gamma_V$, the free VEGF concentration $V$ drops close to zero around the cell-covered area (Fig. 4.19). This change translates into slightly lower overall cell densities, comparable with the effect of an increased inhibitor release rate $\Gamma_*$ (Fig. 4.13) – but does not effect the basic patterning mechanism.
4.5 Discussion

4.5.1 Vascular patterns are formed by multiple guidance systems

Vascular patterning in amniotes is an adaptive, self-organized process. At the tissue scale a hypoxia sensing feedback mechanism controls vascular density [179], and hemodynamic forces of blood flow – that depend on the state of the entire vasculature – can guide the remodeling of vascular network topology [180, 181]. An important element of vascular patterning is the formation of new blood vessels through a process of multicellular sprouting, which is – in itself – a complex process likely involving multiple guidance mechanisms [52].

Endothelial cells can be guided by a pre-pattern in the extracellular environment [125, 126]. Vasculogenesis in fish, where major vessels assemble directly (i.e. without
forming an intermediate vascular plexus) also seems to be guided by a genetic pre-pattern, as specific vascular malformations are correlated with genetic defects [182]. A VEGF pre-pattern, similar to the one in the retina, has not been demonstrated in the ECM associated with the lateral plate mesoderm in avian embryos [183]. Furthermore, endothelial cells can self-assemble vascular networks in artificial culture environment [184, 185], an important factor in recent tissue engineering reports [186].

Endothelial cells could also be guided by ECM structures, which can serve as a spatial “memory” [187–189]: both endocardial [190] and endothelial [191] cells were reported to leave degraded ECM fragments or “channels” behind. Multicellular sprouts readily develop when a positive feedback between the direction of active cell movement (cell polarity) and ECM “memory” is assumed [192–194] – a mechanism very similar to ants use to organize pheromone trails [195]. ECM structures may also be generated by mechanical forces [196–198], a mechanism likely determining patterning on highly malleable matrigel substrates [199].

A special type of cell adhesion, a temporary lateral inhibition mediated through
Delta and Notch cell surface receptors is also operational within angiogenic sprouts [176] and thought to be responsible for restricting the invasive (tip) phenotype to a few cells of the sprout [175]. While lateral inhibition of tip cells through the Delta-Notch system strongly influences vascular morphology, however, branching continued to occur in experiments where Delta-Notch signaling was blocked [200, 201]. Endothelial cells can frequently switch phenotype: leading tip cells are frequently overtaken by follower cells which then assume a tip cell phenotype [176, 201, 202]. The interplay between the Delta-Notch and sVEGFR1 guidance systems remains an interesting problem for future studies.

Endothelial cell movements guided by autocrine chemotactic signaling were proposed as a potential mechanism for vascular pattern emergence [132, 133, 203]. The mechanism relies on the secretion of a diffusing chemotactic morphogen, likely to be VEGF. An autocrine chemoattractant is expected to result in cell aggregation [204], but an interplay between elastic compression and active motility can yield branching patterns [134, 205]. While such autocrine VEGF signaling may also contribute to the patterning of vascular sprouts [133], it is unlikely to be a required mechanism for sprouting activity: endothelial sprouts readily elongate even in the presence of large concentrations of exogenous VEGF in the culture medium [206, 207]. A mathematically similar patterning process can also result from a number of other considered mechanism: If a secreted proteolytic agent increases the availability or “activates” ECM-bound VEGF, then a local gradient of the bio-active VEGF may be produced in the microenvironment of an endothelial cell cluster. Similarly, the binding of paracrine growth factors to angioblast-produced ECM can drive patterning by creating spatially-restricted guidance cues required for directed cell migration [208]. The common feature of these mechanisms is that it operates –contra-intuitively– with attractive gradients pointing towards the endothelial cells.

Following experimental evidence for the possible role of secreted inhibitors in vascular patterning, here we investigate a markedly distinct mechanism, where endothelial cells are guided by self-generated VEGF gradients, pointing away from the vascular structure. Branching patterning under such, diffusion-limited conditions is well studied. In particular, the Mullins-Sekerka instability [209], which renders a smooth interface unstable by triggering a spontaneous tip-splitting process, was shown to yield a characteristic dense branching morphology. As a tip extending into the diffusive external field senses increasingly steep gradients, it further enhances its growth creating the amplification mechanism underlying the instability. Here we explore such a variant of diffusive patterning, in which the gradients are produced by a reaction between an immobile activator (the VEGF) and a diffusive inhibitor (the sVEGFR1). We
demonstrate, that the secretion rate and extracellular stability of the inhibitor can control vascular branching density, but not the expansion speed of branches. How the various endothelial guidance systems are integrated remains an interesting question for future research.

\subsection*{4.5.2 The VEGF system}

VEGF-A and its receptors VEGFR1 (Flt1) and VEGFR2 (Flk1) play a key role in the regulation of vasculogenesis and angiogenesis. VEGF-A isoforms are secreted by a variety of cell types except the majority of the endothelial cells [117, 210, 211], whereas VEGFR1 and VEGFR2 receptors are expressed primarily by those [117, 212]. VEGF-A has multiple isoforms generated by alternative mRNA splicing, having different binding affinities for heparan sulphate proteoglycans (HSPGs) and Neuropilins as well as for VEGF receptors. Except for one, all isoforms contain a HSPG binding domain, and therefore can be sequestered in the extracellular matrix (ECM). Mouse embryos lacking all HSPG-binding VEGF-A isoforms indicated that ECM-binding isoforms are essential for the establishment of steep extracellular VEGF-A gradients [21, 48].

ECM-bound VEGF can be released after a protease cleavage as a diffusible, active protein [20, 21, 112, 213–215]. Recent analysis of available kinetic data, however, do not support the view that a localized VEGF release is required for an effective signalization process – instead it appears to be a multicellular cooperative phenomenon [177]. As the binding sites of VEGF for ECM and sVEGFR1 are distinct, VEGF and sVEGFR1 could interact in the ECM-bound form [216, 217]. Moreover, ECM-bound VEGF (without cleavage) can effectively activate endothelial cells [218, 219]. Given these complexities, in the model we focus on the ECM-bound VEGF population and do not represent the freely diffusing VEGF proteins. Our study suggest, however, that effective VEGF concentrations in tissues are much lower than the 100 pM required in vitro to elicit detectable proliferative or migratory response. This could also reflect the finding, that endothelial cells are more sensitive to VEGF in the presence of appropriate ECM in the tissue environment [125, 157]. Hence, without ECM binding the sVEGFR1-VEGF control system operates in a regime (high $V_*$) which is not sensitive enough to guide patterning.

A variety of VEGF-induced changes (proliferation, migratory activity and chemotaxis) are likely to contribute to the expansion of the vasculature. Vascular sprouts contain highly motile cells, and a directional bias towards the tip provides vascular building blocks in excess of local proliferation both in mouse allantois explants [202] and during the formation of the first vascular plexus in avians [220]. Endothelial cell
streaming along the dorsal aorta also provides cells from the extraembryonic blood islands to the expanding vasculature around the heart [221]. Similar directed endothelial motion is also present in the sprouts forming within a 3D gel environment [201]. Long-range migration of endothelial cells along vascular segments can thus locally expand the vasculature. We condensed this complex process into a phenomenological formalism, directed growth.

In our experiments, we applied an antibody which targets both membrane-bound and diffusive forms of VEGFR1. Hence, while our experiments can test model predictions (Figs. 4.2 vs 4.4), they leave the possibility open that the membrane-bound form of VEGFR1 could also guide endothelial sprouting behavior through some completely different mechanism. For example, tip cell selection involves the VEGF pathway and the excess of tip cells also yields to denser vascular sprouts [200]. The close correspondence of experimental and model simulation results, however, support the plausibility of a vascular patterning mechanism guided by a diffusing inhibitor.

4.5.3 Model choices

The regulation of endothelial cell behavior is in the forefront of research interest due to its relevance in normal development and in various diseases. Recent quantitative models focused on the dynamics of extracellular factors [22], including sVEGFR1 [222]. In this study we focus on how well established molecular interactions can shape the developing vascular morphology. The morphology of the vasculature is of key importance, it determines the density of vascular segments in a given area, thus it determines how blood vessels can supply tissues with oxygen and other necessary factors, or how much tumor cells need to migrate to enter circulation. The aim of this work is to establish how the sVEGFR1/VEGF diffusing inhibitor/activator system can guide vascular patterning. Hence, we intentionally left out several other known details of endothelial cell biology, like the Delta-Notch mechanism underlying the cellular decision to engage in invasive behavior, or the complexity of growth factor processing within cells or in the environment. Instead, we constructed a model which contains the relevant molecular interactions between the morphogens, but is still simple enough to allow mathematical analysis. After establishing the patterning mechanism with the simplest modeling assumptions, we augmented the models to learn how spatial variation in inhibitor production or activator internalization modulate the self-organized patterns.

We used 2D lattice and a phase-field models to represent the interactions of cells and the diffusive inhibitor sVEGFR1. We argue that any mathematical model of a
Chapter 4. Role of soluble VEGFR1 signaling in vascular patterning

complex system contains several hidden assumptions. Therefore, employing multiple and complementary modeling approaches is crucial to increase our confidence in explicitly stated model components. Moreover, both models offer advantages: the lattice model is simple and transparent, while the phase-field method opens up the possibility of analytical calculations. The lattice model can be considered as a variant of the Eden model [223], often used to represent biological growth patterns.

The phase-field method has been successfully applied to various problems, such as dendritic crystal growth [163, 164, 224], morphological changes in biological membranes [225, 226], cell motion [227, 228] and describing the dynamics of multicellular systems [168, 229]. Recently the phase-field approach has been also adapted to model tumor angiogenesis [230–232]. In these models the tumor cells and the adjacent stroma tissue produces VEGF, thereby forming a concentration gradient guiding the growth of pre-existing capillaries. However, during vasculogenesis a steep VEGF concentration gradient is unlikely to develop spontaneously since VEGF is expressed throughout the embryo and a large amount is stored in the ECM. The model proposed here thus expands these previous studies to explain vasculogenesis, the early embryonic vascular patterning process.

Matching model variables to experimentally observable quantities is of key importance in every investigation relying on mathematical models. Here we compare quantities at two scales: both at the level of signaling molecules (like binding and diffusion properties) and at the scale of multicellular structures (like sprout density, branch width and growth speed). Fortunately there are empirical constraints for most of these parameters – albeit their values may vary within an order of magnitude. Our parameters (Table 4.1) includes two unusual choices: the value for VEGFR1 diffusivity ($D_R$), and the value for VEGF threshold ($V_*$). The choice of the low value of $V_*$ reflects our judgment based on the available literature, and it is not required for dense sprouting morphology: simulations performed with $V_* = 100$ pM can also yield branching patterns (data not shown). Two or three orders of magnitude larger diffusivity values, however, were estimated for diffusion of proteins in aqueous solutions or in porous gels [22, 233]. In our model, as well as in other reports [22], simulations performed with such a high diffusivity creates shallow gradients in the order of 1 pM/mm, two orders of magnitude below the experimentally observed VEGF sensitivity threshold [171]. Moreover, $D$ in the range of $10^{-5} – 10^{-4}$ mm$^2$/s yields unrealistically wide branches or compact growth. Such estimates for protein diffusivity, however, do not take into account the binding of sVEGFR1 to the ECM environment. In the presence of reversible adsorption and desorption, theoretical estimates for the effective diffusion parameter scale with the ratio of free and bound molecules [234]. Thus, if the
adsorption of the diffusing agent is efficient enough to keep 99% of the molecules immobilized, the effective diffusivity will drop by two orders of magnitude. Based on the presented data, we suggest that such a limited diffusivity of sVEGFR1 is crucial to guide blood vessel patterning.

4.5.4 Conclusion

The presented results demonstrate that endothelial cells are capable to effectively regulate vascular morphology at the multicellular scale – utilizing a secreted inhibitor and a reaction-diffusion mechanism. The mechanism yields a spontaneous dense branching morphology with a constant vessel density, an ideal distribution architecture for capillaries. In a tissue environment this control mechanism is likely to guide endothelial cells in combination with other factors, like ECM biochemical and mechanical patterning, hypoxia sensing, and various intercellular signaling interactions.
Chapter 5

Cellular cooperativity in tumor recurrence

5.1 Background

The growth of a malignant tumor is driven by the uncontrolled proliferation of cancer cells, and their invasion into healthy tissue. While the primary therapy often involves the surgical removal of the tumor, unfortunately, the surgery often leaves a small population of cancer cells infiltrated into the surrounding tissue. After a remission period of variable duration, the surviving cancer cells can initiate the recurrence of the disease. This is a particularly serious concern for glioblastoma brain tumors characterized by a diffuse tumor boundary within a complex, heterogeneous and relatively soft brain tissue [235, 236].

A major recent retrospective MRI study has shown that 77% of glioma patients relapsed centrally within 2 cm of the original tumor mass, 18% patients relapsed more than 4 cm from the original enhancement and 4% relapsed within the contralateral hemisphere [237]. The median relapse time was 8 month for local relapses, and progressively longer for distant relapses. The median time for contralateral relapses increased almost two-fold, to 15 months.

At the macroscopic level, invasive cancers with a diffuse boundary such as glioblastoma can be described by mathematical models specifying the spatial and temporal changes in tumor cell density [238–243]. Models of tumor invasion often utilize traveling front solutions of the Fisher–Kolmogorov type reaction-diffusion equation [244–246]. Predictive quantitative models of tumor growth have been proposed as a potential tool for patient-specific computational optimization of treatment strategies such as localized radio- and combinatory chemotherapies [247–253]. In combination
with diagnostic imaging, such models aim to forecast the spatial and temporal pro-
gression of the disease taking into account the heterogeneity of the tumor and the
tissue environment [247, 251].

To understand the dynamics that controls the initiation of recurrent tumor growth,
we investigate how surgical removal of the tumor affects its delayed recurrence using
quantitative models. In particular, we aim to identify key parameters of tumor cell
populations that determine how much the progression of cancer can be delayed by
surgical resection. We show that a density dependent proliferation of the cancer cells
[254], particularly at low cell densities, has a key impact on predicting the time until
tumor recurrence.

5.2 Methods

5.2.1 Cell culture

Two human glioblastoma cell lines, U87 and GBM1, were investigated in this study.
U87 is a standard cell line from American Type Culture Collection (ATCC, HTB-14),
GBM1 was established from a giant cell variant of glioblastoma multiforme at the
National Institute of Neurosurgery in Budapest, Hungary as described previously
[255]. Cell lines were maintained and studied in Dulbecco’s Modified Eagle Medium
(DMEM, Lonza) containing L-glutamine, supplemented with 10% fetal bovine serum
(Invitrogen) and penicillin-streptomycin-amphotericin B (Lonza). Cells were grown
in non-precoated culture dishes at 37°C in a humidified, 5% CO₂, 95% air atmo-
sphere. Confluent cultures were washed twice with PBS (Invitrogen) and incubated
with trypsin-EDTA (Sigma) to obtain cell suspensions. Cells were seeded in low den-
sities (3, 10, 30 cells/mm²) into 35 mm Petri dishes (Greiner).

5.2.2 Time lapse imaging

Time-lapse recordings of the cell cultures were performed on a computer-controlled
Leica DM IRB inverted microscope equipped with a Marzhauser SCAN-IM powered
stage and a 10x N-PLAN objective with 0.25 numerical aperture and 17.6 mm working
distance. The microscope was coupled to an Olympus DP70 color CCD camera. Cell
cultures were kept in a stage-top mini incubator at 37°C in humidified 5% CO₂ atmo-
sphere. Phase contrast images were collected every 10 minutes from each microscopic
field for durations up to 3-4 days.
5.2.3 Image analysis

Recorded phase-contrast images were analyzed by segmentation and particle image velocimetry (PIV) algorithms implemented in Octave and Python. To detect cell occupied area a global threshold was applied to the local standard deviation of intensity on each image [256]. The code used for segmentation and confluency calculation are available at http://github.com/aczirok/cellconfluency.

5.3 Results

5.3.1 The Model

We consider a population dynamics model of glioma invasion in which the population density of cancer cells within a tissue is determined by the balance of proliferation, motility and cell death. Tumor cells are known to engage in a rich variety of motility [257]. Yet, as we discuss below, available experimental data suggest that at long time scales cancer cell movement is random and well approximated as a diffusion process, similar to the behavior observed in cell cultures [255]. Thus, tumor spreading at a tissue scale is thought to be well described by a reaction-diffusion equation of the form:

\[
\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left( D \frac{\partial C}{\partial x} \right) + Cr(C),
\]

where \( C(x,t) \) is the density of cancer cells at location \( x \) and time \( t \). The diffusivity of the cells \( D \) characterizes their random motility, and the function \( r(C) \) describes the balance of the rate of proliferation by cell division and cell death rate. In the simplest, and typically used, form of Eq. (5.1) the environment is steady and homogeneous (\( D \) and \( f \) are independent of \( x \) and \( t \)) and the proliferation term is the logistic function

\[
r(C) = \rho \left( 1 - \frac{C}{K} \right),
\]

where \( \rho \) is the maximum population growth rate. Expression (5.2) assumes that, on average, the balance of proliferation and death rates of cells, \( r(C) \), decreases with the cell density and vanishes when the density reaches the carrying capacity \( K \). This behavior reflects – in a simplistic form – limitations of both biochemical resources and cell size as the cell density increases [247, 251, 258, 259].

Equation (5.1) with the logistic proliferation term (Eq. (5.2)) is the well known Fisher–Kolmogorov (FK) equation. The FK equation has traveling front solutions of
the form $C(x, t) = C_0(x - vt)$ where $v$ is the propagation velocity and $C_0$ is the stationary population density profile of cancer cells, as seen in a reference system co-moving with the front [260–262]. For sufficiently localized initial conditions (e.g. with nonzero values restricted to a finite region) the asymptotic front speed is $2\sqrt{D\rho}$ and the characteristic front width is $\sqrt{D/\rho}$.

Following the surgical intervention reactive gliosis appears at the site of surgery. In the majority of the cases for a couple of months the resected area remains tissue free as evidenced by follow-up imaging [263]. As such the cell spreading into this area can substantially be delayed. Thus, in our model it is natural to represent tumor resection (or other localized primary treatments such as radiation therapy) by resetting cancer cell density $C$ to zero in the region where $C$ is higher than a predefined detection threshold $\delta$. Back-propagation of the tumor into the area from which it was removed can be also prevented by no-flux boundary conditions imposed at the contour of the threshold density. The modified cell density profile is then used as initial condition for the same reaction-diffusion equation to generate the post-resection dynamics in the altered spatial domain. We find that our results are quite insensitive to whether or not the resected domain remains available for repopulation.

### 5.3.2 Numerical results: logistic growth

Numerical solutions of Eq. (5.1) with the logistic growth term (Eq. (5.2)) and resection are shown for a one dimensional system in Fig. 5.1. Surprisingly, we find that the resection does not lead to any detectable delay of the propagation of the front: The post-resection front initiated by the truncated, low cell density tail of the cancer cell distribution coincides with the unperturbed original front (see also S1 Movie. B). This behavior appears to be independent of model parameters including the resection threshold $\delta$.

To explain this counterintuitive behavior we note that in the logistic proliferation term (Eq. (5.2)) the cancer-free equilibrium state $C = 0$ corresponding to healthy tissue is linearly unstable. Therefore the FK front is a so called “pulled front” [262, 264], where the dynamics of the low cell density leading edge is not affected even by the complete removal of the population behind the front. The complete absence of a delay in front propagation, however, questions the suitability of FK equation to represent radical medical interventions, which are expected to delay the progression of cancer.
5.3. Results

Fig. 5.1. Space-time plot obtained with the logistic growth term. Numerical solution of the reaction-diffusion (Eq. (5.1)) with the logistic growth term (Eq. (5.2)) in non-dimensional form \((\rho = 1, D = 1, K = 1)\). The initial condition is localized at \(x = 500\) and resection is applied at \(t = 30\) with a threshold \(\delta = 0.1\). After the intervention the simulation is continued using the truncated cancer cell profile as initial condition. (left): The resected region is removed from the domain by imposing no-flux boundary conditions. (right): The density is set to zero within the resected region, but back-propagation remains possible.

5.3.3 Glioblastoma cells exhibit a weak Allee effect in culture

Since after resection the density of cancer cells is low everywhere, the recurrence of the tumor is mainly determined by the survival and proliferation of cancer cells at low cell densities. To gain a qualitative insight into the density dependence of the cancer cell proliferation rate, we performed a series of in vitro experiments.

Glioblastoma cells were grown and imaged in sparse cultures for at least 4 days. Cultures were seeded at low cell densities ranging from 3 to 100 cells/mm\(^2\), corresponding to an area confluency (coverage) between 0.5 and 20\% (see Movies S3 B and S4 B). We evaluated a time-lapse image sequence \(s\) in terms of \(A_s(t)\), the total area covered by the cells as a function of time \(t\) (Fig. 5.2, see Methods for further details).

The growth rate \(b_s(T)\) characterizing the time period \(T \leq t \leq T + \Delta T\) was obtained as a linear fit of the corresponding \(A_s(t)\) values:

\[
A_s(t) = b_s(T)t + \text{const.} \tag{5.3}
\]
Chapter 5. Cellular cooperativity in tumor recurrence

Fig. 5.2. Allee effect in glioblastoma cell cultures. A): U87 glioblastoma cells in culture, detected by our segmentation algorithm (blue). B): Confluency (area covered by cells) as a function of culture time $t$ in parallel cultures of the U87 cell line seeded at various densities, specified in the plot. Each line represents an average of $n = 12$ microscopic fields of the same culture. The standard error of the mean is indicated by the shaded areas. The dashed lines are linear fits over two 30 h long time periods, each starting at a confluency value of 20%. C, D): Average growth rates $r$ for various confluency values $A$, for two glioblastoma cell lines, U87 (C) and GBM1 (D). To obtain each value $r$, we pooled data from three distinct experiments, each monitoring 4 parallel culture dishes seeded at various densities. Thus, each $r$ value incorporates data from 40-100 microscopic fields.

We have chosen the duration of the time period as $\Delta T = 30$ h, sufficiently short for the linear approximation (Eq. (5.3)) to hold, and sufficiently long to detect slow changes in the area covered by cells (Fig. 5.2A).

The density-dependent average growth rate per cell, $r(C) = f(C)/C$, was obtained as

$$r(A) = \langle b_s(T)/A(T) \rangle_{s,t:A_s(T)\approx A'}$$

(5.4)

where the $\langle \ldots \rangle$ average was calculated over parallel cultures $s$ and time intervals for which the initial $A_s(T)$ coverage was sufficiently close to $A$. 
Experimental results from two glioblastoma cell lines suggest that in the low density regime the cell growth rate $r(C)$ increases with the population density $C$ while it decreases at larger densities (Fig. 5.2). This non-monotonous behavior is in contrast with the logistic model which assumes a monotonously decreasing growth rate. In ecology such behavior is known as the Allee effect [265, 266], and can arise as a result of some sort of cooperative behavior among individuals that becomes less efficient at low population density. In cultures of cancer cells such cooperative behavior can likely arise due to autocrine growth factors, diffusive signaling molecules produced and secreted by cells that enhance growth and proliferation of other cells [267]. Mathematical and computational models of cellular mechanisms leading to the development of Allee effect in the context of tumor growth has been described in recent studies [268, 269], and properties of traveling front solutions in a model of tumor invasion with strong Allee effect was studied in [270].

### 5.3.4 Computational model with Allee effect

Motivated by our experimental observations of non-monotonous density-dependent survival and proliferation of tumor cells, we replace the logistic growth rate (Eq. (5.2)) with a quadratic net cell proliferation rate

$$r(C) = \rho \left( \frac{C}{K} + \beta \right) \left( 1 - \frac{C}{K} \right).$$  \hspace{1cm} (5.5)

We choose this functional form as being the simplest that describes a non-monotonous density dependent proliferation. The Allee effect can be categorized by the sign of the parameter $\beta$ as “strong” when $\beta < 0$, or “weak” otherwise. In the case of strong Allee effect the spatially uniform population dynamics is bistable and there is a critical density $C_T = -\beta K$ below which the growth rate is negative. For $\beta > 0$, the case of weak Allee effect, the cell reproduction rate increases with cell density, but it is always positive and there is no critical survival density.

The existence of a minimal density required for the survival of cancer cells would imply that the tumor can be eliminated completely if the resection threshold is sufficiently low ($\delta < -\beta K$). This is, however, very rare in the case of glioblastoma [271] – suggesting that this disease exhibits a weak Allee effect: $0 \leq \beta \ll 1$.

We used the cell proliferation function (Eq. (5.5)) and repeated the tumor growth and resection simulations in one dimension (Fig. 5.3). According to our expectations, in the modified model the resection can indeed substantially delay the propagation
Chapter 5. Cellular cooperativity in tumor recurrence

Fig. 5.3. Space-time plot obtained with Allee effect. Numerical solution of the reaction-diffusion (Eq. (5.1)) including the Allee effect (Eq. (5.5)), in non-dimensional form $\rho = 1, D = 1, K = 1, \beta = 0, \delta = 0.1$.

Fig. 5.4. Total number of cancer cells as a function of time for different values of the resection threshold $\delta$ ($\rho = 1, D = 1, K = 1, \beta = 0$). A rough estimate of the time unit based on the experimentally observed in vitro proliferations rates is about 2 − 3 days. Note, that the case $\beta = 0$, shown in these simulations, gives an upper limit of the delay time, for nonzero $\beta$ the values can be significantly lower.

of the tumor (see also S2 Movie B). Fig. 5.4 shows the integral of cancer cell population density $C(x, t)$ within the area outside the resection, for different values of the resection threshold $\delta$. Note, that after resection there is a lag phase during which the
Fig. 5.5. Remission time obtained with Allee effect. The remission time $\tau$ caused by resecting the tumor as a function of the resection threshold $\delta$ at different values of the parameter $\beta$.

total number of cancer cells is almost constant. The lag phase is followed by a sharp transition to a linear increase indicating a front moving with constant speed. From this graph we can determine the length of the remission period, $\tau$, as a delay relative to the original unperturbed front. The remission period thus increases substantially as the resection threshold is reduced.

The dependence of the remission period length $\tau$ on the resection threshold $\delta$ (Fig. 5.5) is qualitatively different depending on the type of Allee effect considered. In the case of strong Allee effect, the delay time becomes infinite at a finite critical resection threshold. In the borderline case when $\beta = 0$ we find a power law behavior where the delay is inversely proportional to the square of the resection threshold. For weak Allee effect with $\beta > 0$ the remission period length appears to follow a power law similar to the $\beta = 0$ case for larger values of the threshold $\delta$, and crosses into a logarithmic function when the resection threshold is low.

5.4 Discussion

We have shown that the model of tumor invasion based on logistic cell proliferation cannot describe the delayed progression of cancer due to resection and therefore it
Chapter 5. Cellular cooperativity in tumor recurrence

may not describe correctly the typical outcome of clinical interventions that substantially reduce population density of tumor cells. We propose that the key element, that determines the time until tumor recurrence, is the Allee effect, which results from positive cooperative behavior of the cancer cells.

The Allee effect at the level of a tumor cell population may reflect diverse processes at the cellular level. A number of signaling pathways that include autocrine components, such as TGFα/EGF/EGFR, PDGF/PDGFR, HGF/SF and CXCL12/CXCR4 ligand/receptor systems, have been identified in glioma and glioblastoma [267]. Thus, glioblastoma cells can both produce the diffusive factor and respond to its presence with the appropriate receptors that activate cell proliferation. In addition, interactions between tumor cells and the surrounding stromal cells may also depend on the concentration of growth supportive paracrine factors and thus on the local cell density [272]. Furthermore, the matrix remodeling capacity, including the deposition of fibrillar collagen that promote glioma cell invasion, is also influenced by the density of tumor cells [273]. Finally, multicellular spheroid models of tumor growth often exhibit resistance against various treatment modalities [274].

Our mathematical model that includes the Allee effect provides the following insight into the dynamics of the tumor cell population: After resection the proliferation of cancer cells is very slow therefore their distribution is mainly determined by random motility which spreads the cells into the low density regions faster than they could reproduce leading to progressively lower densities. The process is eventually halted by the density distribution of the cells near the resection boundary becoming almost uniform in space. Without a cell density gradient random cell motility cannot further reduce cell density and the slow proliferation eventually catches up and leads to the recurrence of the invasion front. In accord with this analysis, the radiologically and histologically assessed cellularity, i.e. the density of tumor cells in the tissue, is one of the most important histological prognostic factors in glioblastoma multiforme – more predictive than the total tumor burden or proliferation index of the surgical specimen [275–277]. A counterintuitive prediction of our analysis is that reduced cell motility would promote an earlier local recurrence of the disease. Experimentally, this hypothesis could be tested by comparing the migratory activity of patient derived glioma cells and the time of recurrence using a major glioblastoma cohort in order to decrease the impact of other potential confounding factors like genetic background or extent of resection.

Glioblastoma cells are known to follow extracellular matrix rich structures, myelinated tracks and tissue inhomogeneities such as blood vessels or white matter tracts (axon bundles). However, only 20-30% of glioma recurrence is non-local (occurs at
a distance greater than 2 cm from the original tumor centroid) [237, 278]. Thus, remissions clearly involving directed cell migration in great excess to local diffusivity happen, but our simple model representing cell motility as a diffusive process deals with the majority of cases. While to the best of our knowledge there is no single tracking data available for glioblastoma cells in situ or in brain slice explants, in the latter experimental model cells often spread in a spatially isotropic pattern that appears to be consistent with a diffusive spreading [279, 280]. Thus, while glioma motion may be anisotropic and directed at sub-millimeter scales, the complexity of the brain tissue may result in an approximate diffusive spreading at larger scales.

The diffusion term of Eq. (5.1) may also incorporate density-dependent effects. The random motility of cancer cells may also depend on the local cell density, hence affecting the diffusion parameter $D$. When $D$ vanishes for low population densities, the diffusive FK fronts are replaced by compact fronts with a well defined boundary [281, 282]. Similarly, expansion of an adhesive tumor mass without substantial random motility would be described by an advection term. Although, such generalizations are likely to be relevant for other malignancies, the diffuse infiltrate characteristics of glioblastoma are best explained by a diffusive process with a finite $D$ at vanishing densities.

Recent improvements in imaging technology offer the promise of treatments specifically optimized both for the individual patients and tumors at the specific locations. We demonstrated that predictive models of tumor progression, necessary to evaluate and design such treatments, must include the Allee effect of tumor cell population dynamics. We considered a highly simplified one-dimensional model. In reality the strongly non-uniform tissue environment distorts the shape of the tumor and influences the cell’s ability to migrate. Although this will not modify the main qualitative observations regarding the relationship between tumor recurrence time and the Allee effect, such inhomogeneities and tissue anisotropies need to be taken into account when optimizing treatment modalities in a patient-specific manner. While surgery always aims to remove most of the tumor cells, our results indicate that interfering with autocrine feedback regulation of growth control at low cell densities may effectively prolong remission after surgery. As areas with maximal cell densities (and not the total tumor burden) determine remission time, radiotherapy optimization must also critically depend on the Allee effect.
Summary

Elucidation of complex regulatory dynamics underlying collective cell behavior and tissue function is a major challenge of systems biology, that requires a combination of experimental and mathematical or computational methods. In this study, we examined three related problems, where collective behavior emerges through feedback regulatory mechanisms.

Pluripotent stem cell populations exhibit heterogeneous expression levels of Nanog and several downstream transcription factors, however, the underlying molecular mechanism is not yet elucidated. By computational modeling, we observed that instead of negative feedbacks within the core transcriptional regulatory circuit, autocrine FGF signals that act through the MAP kinase cascade are more likely to generate distinct sub-states within the “ON” state of the core NANOG switch. Thus, the experimentally observed fluctuations in Nanog transcription levels are best explained as noise-induced transitions between autocrine feedback-generated sub-states. We also demonstrated that ERK phosphorylation is altered and being anti-correlated with fluctuating Nanog expression, in accord with model simulations.

Vascular patterning is a key process during development and disease, yet it is little understood how the multicellular vascular patterns can be regulated by signaling processes at the molecular or subcellular level. By employing established computational models of VEGF signaling, we investigated how a molecular inhibitory autocrine signaling system can regulate the structure of growing vascular sprouts. We demonstrated that a diffusive inhibitor (sVEGFR1) can generate structures with a dense branching morphology. Inadequate presence of the inhibitor leads to compact growth, while its excessive production blocks expansion and stabilizes existing structures. Model predictions were validated with time-resolved experimental data obtained from endothelial sprout kinetics in fibrin gels.

Resection of the bulk of a solid tumor, such as glioblastoma, often cannot eliminate all cancer cells due to their infiltration into the surrounding healthy tissue. By employing a generalized Fisher-Kolmogorov equation, we investigated how the invasion front is delayed by resection, and how this depends on the density and behavior of the remaining cancer cells. While the typically assumed logistic proliferation term leads to unrealistic recurrence results, the introduction of an experimentally established cooperative behavior of cancer cells (Allee effect), seems to be critical in determining the time until tumor recurrence.
Összefoglalás

A sejtek kollektív viselkedésének és így a szöveti működést meghatározó szabályozás megértése a rendszerbiológia (systems biology) egyik fő kihívása. A rendszer komplexitása miatt nem csak kísérleti vizsgálatokra, hanem matematikai és számítógépes modellek felállítására is szükség van. Ebben a munkában három olyan problémát vizsgáltunk, melyekben a sejtek kollektív viselkedése visszacsatoláson alapuló szabályozási mechanizmusokon keresztül valósul meg.

Pluripotens össejt populációkban a NANOG, és számos a NANOG által vezérelt (downstream) transzkripciós faktor expressziós szintje heterogén. Számítógépes modellezéssel megmutattuk, hogy a NANOG rendszer "BE" állapotán belül különböző alálaplatokat hozhat létre az autokrin FGF-MAPK jelátvitel, míg ez nem valószínűsíthető a transzkripciós szabályozó modul negatív visszacsatolásai esetén. A \( \text{Nanog} \)
transzkripciós szintjeiben tapasztalt inhomogenitás így leginkább zaj-indukált átmenetekként magyarázhatóak egy, az autokrin rendszer által létrehozott alálaplatok között. Célzott kísérletekkel megmutattuk, hogy a szimulációs jósolatokkal összhangban egér embrionális össejtben az ERK foszforiláció változó, és antikorrelált a NANOG menyanyiségével.

Az erek kialakulása az embrionális fejlődés és számos betegség meghatározó folyamata, a többejtű érintészetek molekuláris és szubcelluláris szintű jelátviteli folyamatokkal való szabályozása kevésbé ismert. A VEGF jelátvitelt leíró számítógépes modellek alkalmazásával megvizsgáltuk, hogy egy molekuláris gátlásan alapuló autokrin jelátviteli rendszer hozzák létre a növekvő ércsírákhoz a \( \text{Vegfr} \)
transzkripciós szintjeiben tapasztalt inhomogenitás és negatív visszacsatolások. Megmutattuk, hogy egy diffúzív inhibitor (sVEGFR1) sűrű, elágazó morfológiájú (dense branching) struktúrákat hozhat létre. Az inhibitor korlátozott jelenléte kompakt növekedéshez vezet, míg túlzott termelése stabilizálja a már meglévő struktúrát, gátolva azok további terjedését. A modell jósolatai fibrin gélben növekvő endotél csírák invázióját leíró kísérleti adatokkal hasonlítottuk össze.

Egy szolid tumor sebészeti kimetszésével gyakran nem távolítható el az összes rákos sejt, mivel azok a környező, egészséges szövetekben is megtalálhatóak. Ez az invazív viselkedés különösen fontos egy rosszindulatú agydaganat, a glioblasztómá esetén. Kísérletekkel megmutattuk, hogy glioblasztómá sejtek proliferációjá és mozgása függ a sejszűréségtől. Egy általánosított Fisher-Kolmogorov egyenlet alkalmazásával megmutattuk, hogy ez a tumorjelek közötti kooperatív jelenség (Allee effektus) meghatározó jelentőségű a tumor remissziós idejének meghatározásában.
Acknowledgements

First of all, I would like to thank my supervisor, András Czirók, for his guidance and enormous support throughout my PhD studies, and also for introducing me to various interesting biophysical and cellular biology topics. I feel grateful for the knowledge and mentality, and also for all the computational methods he passed on to me.

Special thanks go to Júlia Táróski-Zách, Előd Méhes and Márton Gulyás, who have supported me with their immense experience in cell biology and time-lapse microscopy. I also would like to thank the support of the people at the Biological Physics Department. I am also very grateful for the members of the Oncology Lab at the Medical University of Vienna, especially for Viktória László, Balázs Döme and Balázs Hegedűs for the opportunity to work there and learn new experimental techniques.

I am enormously grateful to Tomi for supporting me all the way, particularly for the Csepel-style assistance throughout the years. I would like to express my very great appreciation to my friends from university, especially the residents of Bartók32, the retired Mafihe members and the community of Láncostáplálék, whom without this thesis would not have been possible. I am also thankful to my classmates and Pokol fellows for providing the upbeat in various ways all the time. My special thanks are extended to Máni and Ákos for their help in the maintenance of my physiological state. Furthermore, I would like to thank my family and Bagira for their support.
Appendix A

Chapter 4 - supplementary movies

Supplementary movie

Sprouting assay using HUVEC spheroids in fibrin gel in absence or presence of VEGFR1 antibodies. Movie composed of time-lapse phase-contrast image sequences of a parallel untreated control culture (left), and a culture exposed to function blocking antibodies against VEGFR1 at a concentration of 20µg/ml (right).

https://osf.io/njd7u
Appendix B

Chapter 5 - supplementary movies

Supplementary movie 1

Movie showing the front propagation before and after resection using the logistic proliferation term. Blue curve is the original unperturbed front and Red dashed curve is the post-resection front. Note, that in this case the post-resection front coincides with the unperturbed one, i.e. the front propagates without any delay. The parameters are $D = 1, \rho = 1, K = 1$, the resection time is $t_s = 20$ and resection threshold is $\delta = 0.1$.

https://doi.org/10.1371/journal.pcbi.1005818.s001

Supplementary movie 2

Movie. Same as S1. except that the proliferation term is replaced with the cubic function including the Allee effect ($\beta = 0$). In this case the resection is followed by a latent remission period and the recurrent front is substantially delayed relative to the original front. The parameters are $D = 1, \rho = 1, K = 1$, the resection time is $t_s = 70$ and the resection threshold is $\delta = 0.1$.

https://doi.org/10.1371/journal.pcbi.1005818.s002

Supplementary movie 3

Movie. Segmented time-lapse phase-contrast image sequence of the U87 cell line.

https://doi.org/10.1371/journal.pcbi.1005818.s003

Supplementary movie 4

Movie. Segmented time-lapse phase-contrast image sequence of the GBM1 cell line.

https://doi.org/10.1371/journal.pcbi.1005818.s004
Bibliography


Bibliography


ADATLAP
a doktori értekezés nyilvánosságra hozatalához*

I. A doktori értekezés adatai
A szerző neve: Lakatos Dóra
MTMT-azonosító: 10048275
A doktori értekezés címe és alcíme: Collective features in the control of cell behavior
DOI-azonosító: 10.15476/ELTE.2018.069
A doktori iskola neve: Fizika Doktori Iskola
A doktori iskolán belüli doktori program neve: Statisztikus Fizika, Biológiai Fizika és Kvantumrendszerek Fizikája program
A témavezető neve és tudományos fokozata: Dr. Czirók András, PhD
A témavezető munkahelye: Eötvös Loránd Tudományegyetem, Biológiai Fizika Tanszék

II. Nyilatkozatok
1. A doktori értekezés szerzőjeként
a) hozzájárulok, hogy a doktori fokozat megszerzését követően a doktori értekezésem és a tézisnek nyilvánosságra kerüljenek az ELTE Digitális Intézményi Tudásvárosban. Felhatalmazom a Természettudományi kar, Dékanál, Hivatali Doktori, Habilitációs és Nemzetközi Ügyek Csoportjának ügyintézőjét, hogy az értekezést és a téziseket feltöltsé az ELTE Digitális Intézményi Tudásvárosba, és ennek során kitöltse a feltöltséhez szükséges nyilatkozatokat.

b) kérem, hogy a mellékelt kérelmemben részletezett szabadalmi, illetőleg oltalmi bejelentés közzétételéig a doktori értekezést ne bocsássák nyilvánosságára az Egyetemi Könyvtáran és az ELTE Digitális Intézményi Tudásvárosban;

c) kérem, hogy a nemzeti biztonsági okból minősített adatot tartalmazó doktori értekezést a minősítés (datum)-ig tartó időtartama alatt ne bocsássák nyilvánosságára az Egyetemi Könyvtáran és az ELTE Digitális Intézményi Tudásvárosban;

d) kérem, hogy a mű kiadására vonatkozó mellékelt kiadó szerződésre tekintettel a doktori értekezést a könyv megjelenéséig ne bocsássák nyilvánosságára az Egyetemi Könyvtában, és az ELTE Digitális Intézményi Tudásvárosban csak a könyv bibliográfiai adatait tegyék közzé. Ha a könyv a fokozatkereszt követően egy évig nem jelenik meg, hozzájárulok, hogy a doktori értekezésem és a tézisnek nyilvánosságra kerüljenek az Egyetemi Könyvtában és az ELTE Digitális Intézményi Tudásvárosban.

2. A doktori értekezés szerzőjeként kijelentem, hogy
a) az ELTE Digitális Intézményi Tudásvárosba feltöltendő doktori értekezés és a tézisek saját eredeti, önálló szellemi munkám és legjobb tudomásom szerint nem sértet vele senki szerzői jogait;

b) a doktori értekezés és a tézisek nyomtatott változatai és az elektronikus adathordozóban benyújtott tartalmak (szöveg és ábrák) mindenben megegyeznek.

3. A doktori értekezés szerzőjeként hozzájárulok a doktori értekezés és a tézisek szövegének plágiumkereső adatbázisba helyezéséhez és plágiumellenőrző vizsgálatok lefuttatásához.


*ELTE SZMSZ SZMR 12. sz. melléklet