

UNIVERSITÉ PARIS 7

École Doctorale « Frontières du Vivant »

T H È S E

pour obtenir le titre de

Docteur de l'Université Paris Diderot (Paris 7)

Mention : FRONTIÈRES DU VIVANT

présentée par

Julien DELILE

**Des comportements cellulaires à la morphogenèse
embryonnaire : modélisation mécano-génétique et
simulations computationnelles du développement
animal précoce**

Thèse dirigée par Nadine PEYRIÉRAS et René DOURSAT

préparée à

l'Institut des Systèmes Complexes Paris Île-de-France (ISC-PIF)
et au Laboratoire Neurobiologie et Développement (NED), CNRS Gif-sur-Yvette

soutenance le 14 Février 2013

<i>Rapporteurs :</i>	Lance DAVIDSON	- Université de Pittsburg (USA)
	Roeland MERKS	- NISB - CWI (Pays-Bas)
<i>Examinateurs :</i>	Dirk DRASDO	- INRIA Rocquencourt
	Ed MUNRO	- Université de Chicago (USA)
	Liliana SOLNICA-KREZEL	- Université de Saint Louis (USA)
	Sophie VRIZ	- Université Paris Diderot
<i>Directeurs :</i>	Nadine PEYRIÉRAS	- NED - CNRS (UPR 3294)
	René DOURSAT	- Université Drexel (USA)

Acknowledgements

First of all, my gratitude goes to my PhD supervisors, *Nadine Peyriéras* and *René Doursat*, who shaped my vision of what being a researcher requires and means. Their association was complementary, Nadine communicating the flame and energy necessary to undertake this vast project, and René bringing his crystal-clear precision to formalize the model; Nadine opening her lab and advising me on the biological concepts and challenges, and René overseeing their abstraction into equations and agent-based simulations. I am especially grateful to both of them for the precious help that they provided in the final months of the project, leading to the fulfillment of the present manuscript and concluding these four years of work.

I am very grateful to *Liliana Solnica-Krezel*, *Sophie Vriz*, *Dirk Drasdo*, *Ed Munro*, *Lance Davidson* and *Roeland Merks* to honor my defence jury with their presence, and particularly to *Lance Davidson* and *Roeland Merks* for accepting to review the manuscript and provide insightful comments, despite a somewhat stretched out delivery on my part.

I also want to thank my PhD tutor *Stéphane Doncieux* for his wise pieces of advice, particularly concerning parameter space exploration strategies, and to *Pierre Collet* for his collaboration to this project and for allowing the simulation of thousands of virtual zebrafish embryos in the *j4 room* of the university of Strasbourg.

Naturally, this project could not have come into existence without the Bourse de Docteur Ingénieur (BDI) scholarship, provided by the French National Center for Scientific Research (CNRS) and the General Directorate for Armament (DGA). I also wanted to thank my PhD program, Frontiers in Life Sciences, under the direction of *FranTaddei* and *Samuel Bottani*, for providing such a inspiring environment, ideal for the realization of innovative projects.

Creativity and interdisciplinarity are two key words describing the place where I have been fortunate to spend most of these four years in the center of Paris, i.e. the Complex Systems Institute, Paris Île-de-France (ISC-PIF). I am very thankful to its founding director, *Paul Bourgine*, whose visionary and inexhaustible initiatives are the *cornucopia* from which so many interdisciplinary “complex systems” projects were born, including this unique environment. I am also thankful to his successors, *René Doursat*, *Arnaud Banos* and *David Chavalarias*, for pursuing the legacy and developing new opportunities of intellectual exchanges. The “ISC” has hosted and continues to host great personalities whose continuous interactions with me have greatly contributed to both the scientific content of this PhD thesis and the daily enjoyment of these four years. I may cite *Daniel*, *Romain*, *Pierre*, *Mathieu*, *Matthieu*, *Mathieu*, *Tam Kien*, *Barbara*, *Thierry*, *Emmanuel*, *Camilo*, *Benoît*, *Raphaël*, *Carlos*, *Elisa*, *Marlene*, *Geneviève*, *Stéphanie*, *Jean-Philippe*, *Ricardo*, *Bivas*, *Taras*, *Alberto*, *Davide*, *Julie*, *Julie*, *Julian*, *Elias* and many others (some of them may be found at

<http://iscpif.fr/researchers>). I know this rough list does not transcribe the diversity and the quality of the exchanges we had but a few personalized words would not do them justice anyway.

In parallel, the members of the Neurobiology and Development lab in Gif-sur-Yvette, under the direction of *Philippe Vernier*, have guided the first and last efforts to build this project, from the early fibronectin cloning attempts to the final manuscript redaction. Many thanks to *Gaelle, Adeline, Louise, Paul, Dimitri, Ludovic*, the Colin sisters *Ingrid* and *Audrey, Monique, Sophie, Benoît, Odile*, and the mysterious madeleine donor, for their pleasing company and support.

Last but not least, my warm thoughts go to my friends from Bordeaux, Nancy and Paris, and my family. I want to apologize for having neglected them so often during those four years and I promise to make up for the lost time in the near future. Finally, thanks to *Maria...* for sharing your time with me.

Résumé

Ce travail présente un modèle théorique de morphogenèse animale, sous la forme d'un système complexe émergeant de nombreux comportements cellulaires. Son implémentation repose sur un système multi-agents fondé sur le couplage mécano-génétique entre les dynamiques de la mécanique cellulaire et de la régulation génétique et moléculaire. Notre objectif est l'intégration des mouvements collectifs de myriades de cellules avec les dynamiques d'expression génétique sous-jacentes aux motifs des champs morphogénétiques. Nous examinons aussi les relations de causalité ascendante reliant les comportements cellulaires locaux aux déformations tissulaires globales. Le comportement mécanique de chaque cellule est associé à leur constitution moléculaire et génétique. Nous nous intéressons particulièrement au phénomène d'intercalation cellulaire induit par des comportements de protrusion active. Différents épisodes morphogénétiques se déroulant au cours des 10 premières heures de développement du poisson zébré sont explorés: la période de clivage, la formation de la couche épithéliale externe, l'épibolie, l'internalisation du mésendoderme et la convergence-extension de l'axe antéro-postérieur. Pour chacun de ces phénomènes, une étude de cas examine le rôle respectif des différents tissus impliqués. Les hypothèses que nous proposons sont discutées au moyen de comparaisons automatisées entre les mesures reconstruites issues des données d'imagerie microscopique et des simulations computationnelles. Les reconstructions présentées incluent les champs de déplacements cellulaires *in toto* en 4 dimensions (3D + temps) de l'embryon de poisson zébré. Le modèle est paramétré à travers une interface inspirée du concept de paysage épigénétique de Waddington, permettant de spécifier les phénomènes d'induction et de d'interaction mécanique à l'échelle des champs morphogénétiques.

Abstract

We present a theoretical model of animal morphogenesis construed as a self-organized phenomenon emerging from a complex system made of a myriad of individual cell behaviors. It is implemented in an agent-based simulation centered on the mechanico-chemical coupling between cellular and genetic dynamics. The goal is to integrate the collective motion of cells and the dynamics of their gene expression underlying the patterning of morphogenetic fields. We also investigate the causal bottom-up link from local cell behavior to global tissue deformation. Each cell's mechanical behavior is mapped from its molecular and genetic identity. Among these behaviors, we focus particularly on cell intercalation as an active process driving tissue deformation. We operate this model to explore the different morphogenetic episode occurring through the first 10 hours of the zebrafish development: cell segmentation, enveloping layer formation, epiboly, internalization and convergence-extension. For each specific episode, a case study is realized to decipher the respective role of the different tissue involved. Quantitative measures reconstructed from both the simulated and the experimental data are compared to automatically explore the multi-dimensional parameter spaces of our hypotheses and their interpretation. Various state of the art computational reconstruction will be presented, including global 4D (3D + time) displacement fields from *in toto* data of the developing zebrafish embryos. A waddingtonian interactive timeline tool to specify intra and inter tissue induction and mechanical behaviors is also proposed.

Contents

1	Introduction	13
1.1	A Historical Timeline of Developmental Biology	13
1.1.1	The Science of How Organisms Form	13
1.1.2	First Theories: Epigenesis, Preformation and Spontaneous Generation	14
1.1.3	The Rise of Experimental Embryology	16
1.1.4	Developmental Genetics	19
1.1.5	Molecular Genetics	20
1.1.6	Cell Biomechanics	23
1.2	Integrating Developmental Mechanics and Genetics: The MECAGEN Project	24
1.3	Methodological Considerations	26
1.3.1	The Scientist in the Observation-Hypothesis-Experiment Loop	27
1.3.2	Designing Tools to Perceive, Conceive and Manipulate	29
1.3.3	Reality Check: Validating the Hypotheses	39
1.4	Overview of this Dissertation	42
Part I	A GENERIC MODEL OF MULTICELLULAR DEVELOPMENT	47
2	State of the Art and Principles of Embryo Development Models	49
2.1	A Review of Theoretical Models of Development	49
2.1.1	Early Attempts	49
2.1.2	Reaction-Diffusion Systems	50
2.1.3	Morphogen Gradients and Positional Information	51
2.1.4	Epithelial Cell Shaping and Division Patterns	54
2.1.5	Differential Adhesion and Cell Sorting	55
2.2	Common Modeling Principles: Toward an Integrated Theory of Development	57
2.2.1	Dealing with Multiple Levels of Organization	58

2.2.2	Relating One Level to the Next: The Problem of Emergence	61
2.2.3	Identifying Custom Laws at Each Level	63
2.2.4	Summary Table	64
2.3	Overview of the MECAGEN Modeling Principles	67
3	MECAGEN Model of Cell Biomechanics	69
3.1	State of the Art in Biomechanical Modeling	70
3.1.1	The Many Viewpoints on “Soft Matter”: From Continuous-Space to Discrete-Element Models	70
3.1.2	Particle-Based Physics	82
3.2	Hypotheses	89
3.2.1	Generic Cell-Centered Equation of Motion	89
3.2.2	Cell Neighborhood	92
3.2.3	Interaction Potential	99
3.2.4	Cell Specific Behavioral Forces	106
4	MECAGEN Model of Molecular and Genetic Regulation and Signaling	113
4.1	Intra-Cellular Gene/Protein Reactions	116
4.1.1	Protein-Protein Interactions	116
4.1.2	Synthesis by Encoding Genes	117
4.1.3	Degradation by the Molecular Environment	117
4.1.4	Cis-Regulatory Boolean Model of Gene Expression	118
4.2	Signal Secretion and Transduction Modules	118
4.2.1	Signal Secretion Module	118
4.2.2	Signal Transduction Module	119
4.3	Extra-Cellular Reactions, Transport and Diffusion	120
4.4	Illustration on Artificial GRN Motifs	121
5	MECAGEN Model of Mechanic-Genetic Coupling	125
5.1	Toward a Cell Behavior Ontology	126
5.1.1	Cell Behavior Ontology	127
5.1.2	Mechanotransduction Input of the GRN	128
5.1.3	Cell Adhesion	128
5.1.4	Cell Polarization	129
5.2	Waddingtonian Timeline Specification	131
5.2.1	Cell Types	132
5.2.2	Ligand Sinks and Sources	135
5.2.3	Passive Force’s Adhesion Modulation	136
5.2.4	Active Cell Behavior Specification	136
5.3	Illustration on Artificial Cell Sorting	138
5.3.1	Cell Sorting: Revisiting a Classical Problem	138

5.3.2 Individual Cell Migration	145
Part II STUDY OF THE ZEBRAFISH EARLY EMBRYOGENESIS	149
6 Review of the Zebrafish Early Developmental Mechanics	151
6.1 Cleavage Stages, Formation of the EVL and YSL	152
6.1.1 The Formation of the Enveloping Cell Layer (EVL)	154
6.1.2 The Formation of the Yolk Syncitial Layer (YSL)	154
6.2 Blastula Stages and the Onset of Epiboly	155
6.3 Gastrulation Movements	158
6.3.1 Epiboly Phase Two	158
6.3.2 Hypoblast Formation	162
6.3.3 Convergence-Extension	165
7 Experimental Reconstruction of the Zebrafish Early Development	167
7.1 Bioemergences Reconstruction Workflow	168
7.1.1 Embryo Preparation and Acquisition	169
7.1.2 Cell Lineage Reconstruction Workflow	170
7.1.3 Post Processing and Exploitation	174
7.2 Reconstructing In Toto Deformation Fields	175
7.2.1 Manual Registration	177
7.2.2 Voxel Quality Evaluation	181
7.2.3 Blending Function	185
7.2.4 Deformation Fields	186
7.2.5 Blended Deformation Fields	188
8 Simulated Reconstruction and Model Validation Through Case Studies	193
8.1 Investigating the Yolk Biomechanical Properties	195
8.1.1 Hypotheses and Model	196
8.1.2 Simulation, Parameter Space and Validation	202
8.1.3 Discussion	210
8.2 Cell Proliferation Rate Along the Cell Lineage	211
8.2.1 Hypotheses and Model	214
8.2.2 Simulation, Parameter Space and Validation	217
8.2.3 Discussion	227
8.3 Shaping the Zebrafish Blastula	227
8.3.1 Hypotheses and Model	228
8.3.2 Simulation, Parameter Space and Validation	240
8.3.3 Discussion	242

8.4	Cell Behaviors in the Enveloping Cell Layer Compartment	244
8.4.1	Hypotheses and Model	244
8.4.2	Simulation, Parameter Space and Validation	247
8.4.3	Discussion	258
8.5	Intercalation Patterns	262
8.5.1	Hypotheses and Model	263
8.5.2	Simulation, Parameter Space and Validation	269
8.5.3	Discussion	284
8.6	Gastrulation	286
8.6.1	Hypotheses and Model	290
8.6.2	Simulation, Parameter Space and Validation	290
8.6.3	Discussion	291
9	Conclusion	293
9.1	Criteria for a “Good” Model	293
9.1.1	Similarities Between Variables and Concepts	294
9.1.2	Computational Feasibility	295
9.2	Choices Concerning the Biomechanics	296
9.2.1	Spatial Neighborhood	296
9.2.2	Interaction Potential and Behavioral Forces	297
9.2.3	External Objects	299
9.2.4	Single-Cell vs. Subcellular Mechanics	299
9.2.5	Spherical-Voronoi Approximation	299
9.3	Choices Concerning the Molecular/Genetic Regulation and Signaling	300
9.4	Assessment of the Case Studies	301
9.5	Future Work	304
9.5.1	Portability to Other Animal Models	304
9.5.2	Interactive Exploration of Parameter Space	305
9.5.3	Toward an “Evo-Devo” Perspective	305

CHAPTER 1

Introduction

The spontaneous making of an entire multicellular organism from a single cell ranks among the most exquisitely complex phenomena in nature. Through a precise spatiotemporal interplay of genetic switches, chemical signals and mechanical constraints, an elaborate form is created without any of its myriad cells containing the explicit map of the resulting architecture. An eternal source of fascination for generations of philosophers, artists and scientists, biological morphogenesis is the epitome of what can be called today a *self-organizing complex system*. To follow the metaphor used by Enrico Coen in his beautiful popular science book *The Art of Genes*[1], it could be said that the embryo is similar to a “canvas that paints itself” (where colors represent cell differentiation) at the same time that it is growing and “sculpting itself”, too—both patterning and shaping affecting each other in a tight loop. Schematically, the mechanical properties of cells, such as their adhesion strength or intrinsic motility, are tightly correlated with their spatiotemporal location and molecular and genetic dynamics, which define distinct morphogenetic fields that further expand, reshape and segment themselves into subregions due precisely (in a feedback loop) to the self-assembly of differentiating cells.

In this introduction chapter, we first propose a brief chronological review of the field of developmental biology, born from the classical “embryology” (Section 1.1), then state the overall objectives of the project, which is called MECAGEN to highlight the double and coupled mechanical-genetic dynamics of development, to which this dissertation is contributing (Section 1.1.2). This will be followed by general preliminary remarks on the methodological principles and workflow that will be at the foundation of our work (Section 1.1.3), then an overview and summary of the remaining eight chapters of this dissertation (Section 1.1.4).

1.1 A Historical Timeline of Developmental Biology

1.1.1 The Science of How Organisms Form

A general definition would state that *development* is a dynamical process leading a given organism to a certain morphological state, and that the study of development is the study of the mechanisms ruling the coordination of cellular differentiation through space and time in a multicellular organism. This process is not steady, however, and organisms may alternate periods of intense transformation with “quieter” periods dedicated to growth only. However, the most dramatic events occur in

the beginning, when the egg divides into a myriad cells. These soon start to perform a collective ballet of complex movements, which are precisely coordinated through a system of physicochemical interactions. It is interesting to note that this process never ends, the morphological state of an organism undergoing constant change, albeit smaller, until senescence and death.

The definition of development has its own “embryogenesis”: it has also changed and reformed itself through the numerous discoveries and practical methods that have punctuated the history of the field. The following major periods can be distinguished, adapted from Hopwood [2]:

- pre-1880: classical descriptive embryology (mostly addressed in Section 1.1.2)
- 1880-1930: classical experimental embryology (mostly addressed in 1.1.3)
- 1930-1960: reconciling genetics and embryology (mostly addressed in 1.1.4)
- 1960: modern developmental biology (mostly addressed in 1.1.5-6)

1.1.2 First Theories: Epigenesis, Preformation and Spontaneous Generation

Aristotle envisioned them all

If Hippocrates is often labeled as the first embryologist, based on his *Hippocratic Collection* where he describes the successive stages of incubation of hens' eggs [3], it is Aristotle in *De Generatione Animalium* who initiated the main theory of the formation of living organisms that still hold today. This theory, called “epigenesis”, states that organisms develop through changes in shape. New parts appear by themselves and deform through a series of stages. In Book II of his work, Aristotle foresees the alternative theory of “preformation”, which persisted well through the 18th century before it was gradually dismissed. It states that organisms develop from miniature versions of themselves, thus no new structures are truly generated during development, but parts simply unfold and grow:

“Either all the parts, as heart, lung, liver, eye, and all the rest, come into being together or in succession, as is said in the verse ascribed to Orpheus, for there he says that an animal comes into being in the same way as the knitting of a net”[4].

A third hypothesis was developed by Aristotle in *Historia Animalium*[4]. Synthesizing contemporary thought, he introduces the theory of “spontaneous generation”. This theory proposes that some complex organisms such as flies can be generated from non-living matter such as putrefying earth, vegetable matter, or dead flesh.

Early mechanical interpretations

In 1651, William Harvey was among the first philosophers to consider embryonic development as a dynamic process requiring a causal explanation of a sequence of events [5]. In 1664, René Descartes's *De la formation du foetus*[6] was posthumously published and contained the first attempt to explain the formation of an animal embryo by means of mechanistic interactions. In this theory, the male and female seeds were thought to "heat up" and "ferment" until some of their particles started moving and forming the organs, from the heart that pushed the blood directly towards the place that it was freest to go, and participate in the formation of the brain [7].

The end of spontaneous generation

In 1668, Gregorio Redi realized an experiment which discredited the spontaneous generation hypotheses by showing that no fly can form in a closed jar containing food or dead flies. As maggots would appear, however, when the same experiment was run in an open jar, he concluded that the food or dead flies could not produce maggots by themselves [8]. The final blow to this theory was struck by Louis Pasteur two centuries later (in 1859) when, as a young chemist, he showed that a boiled meat broth would not produce new organisms. The key idea of his experimental setup was to bend the neck of the container into a 'S' shape. It allowed air to pass through it but not the aerial microorganisms because of gravity. However, when the inflection point was put in contact with the broth, a culture would rapidly start to grow.

The end of preformationism

The contention between preformists and the upholders of epigenesis lasted longer. Antonie van Leeuwenhoek was a Dutch scientist who created various microscopes. In 1676, he made the first observation of single-celled organisms, "animalcules", soon after Robert Hooke had first described and termed the "cells"[9]. Leeuwenhoek discovered that the sperm cells of animals, among which humans, were entering the egg cell [10]. In addition to his contribution to the refutation of spontaneous generation, this discovery favored the spermist side of the preformation camp. Some of them started to describe miniatured humanoid shapes as did Nicolaas Hartsoeker in 1695 1.1.

Germ layers

In the 1820's, Christian Pander conducted a reinvestigation of developing chicks in egg and explained that development does not start from the formation of organs but originates from the transformation of primitive sheets of tissue, called the "germ layers"[2].

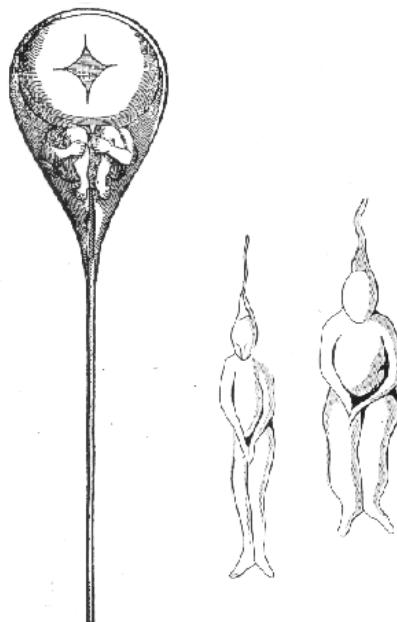


Figure 1.1: Drawing of the homonculi observed in sperm by Nicolaas Hartsoecker (1695)

Cell theory

Between the 1820's and the 1850's, *cells* were added as the second pillar of embryological analysis mostly under the influence of Johannes Müller [2]. In the late 1830's, the "cell theory" attempted to unify the development of the various observed eggs in vertebrates, and particularly in mammals. Cells progressively became the fundamental building blocks of every living species in the minds of the scientists. Robert Remak stated that every cell was produced by a preexisting cell, from the egg all the way to the tissues, via the germ layers [11]. This insight is now called the "segmentation" or "cleavage" stage and is indeed the first morphological event of today's developmental studies. Remak also introduced the concept of germ-layer specificity in vertebrates, stating that each layer (endoderm, mesoderm, and ectoderm) is specifying the cell type or "fate" of all cells that originate from it (such as muscle, skin, nervous system, or intestine) [12]. This concept was also central to embryology and preludes the fundamental questions that will continue occupying developmental biologists of the modern era.

1.1.3 The Rise of Experimental Embryology

Starting in the 1880's, some embryologists reinvented their methods through experimentation to decipher the causal links between the successive stages of development.

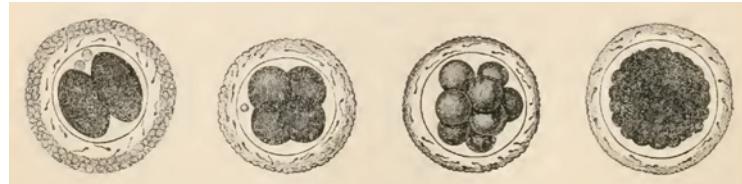


Figure 1.2: Egg segmentation from a dog’s oviduct, surrounded by the zona pellucida and spermatozoa as represented by Albert Kölliker in 1861 [13].

Calling this discipline *Entwicklungsmechanik* (“developmental mechanics”), Wilhem Roux and others applied to embryos various kinds of perturbations, whether mechanical, thermal, chemical or electrical [14].

Entwicklungsmechanik: self-differentiation vs. dependent differentiation

The key question raised by Roux was whether the differentiation process of the parts of an embryo was autonomous from external influence (“autonomous differentiation” or “self-differentiation”) or not (“dependent differentiation”). In 1888, he obtained half-embryos after destroying one of the cells of a two-cell frog embryo with a hot needle. The half-embryos were displaying either the anterior or the lateral halves. Roux concluded that each blastoderm was capable of *self-differentiation*, independently from the missing half [15]. A year earlier, in 1887, Laurent Chabry had been the first to characterize the autonomous differentiation of cells’ fate. By killing two identified blastomeres at the 8-cell stage of the ascidian tunicate, the animal became a tadpole that was missing its tail muscles. When he extracted and cultured the same two blastomeres at the same stage, they resulted in an isolated tail muscle [16].

In 1891, following Roux’s influence, Hans Driesch repeated the experiment on a two-cell stage sea urchin embryo. He separated both blastomeres and observed that each one had differentiated into a half-sized, yet complete, sea urchin larva. In 1893, by pressing on a sea urchin embryo at the third cell cycle, Driesch completely modified the relative positions of the cells and still obtained normal larvae. A mosaic determination process would have produced a highly perturbed embryo, therefore it proved that the determination occurred later than expected via *dependent differentiation*. Driesch concluded: “The relative position of a blastomere within the whole will probably in a general way determine what shall come from it.”

These historical experiments epitomized the concurrent interpretations of autonomous vs. dependent differentiation, where the former requires “determinants” to be present at the earliest stage and separated by cell division to spatially specify cell fates, whereas the latter depends on the interaction between the cells (see a classification in [17]). Later, it was recognized that truth resided in the middle. Neither

totally mosaic nor totally regulative, developmental principles are a fine balance between both principles [16]. Some cells at certain stages seem highly dependent from their surroundings, and other times they seem to “seal their hatches” and follow their own differentiation path. From there on, most embryological studies will be dedicated to deciphering the modus operandi of these mixed principles, from their macroscopic characterization at the tissue level down to the molecular mechanisms at the sub-cellular level.

Morphogenetic fields

The early 20th embryologists refined these questions with new experiments such as grafts and aimed at deciphering what determined cell fates. In 1918, Ross Harrison published a paper that introduced the concepts of *morphogenetic fields*[18]. He carried out various limb grafts on the newt embryo. He transplanted some cells from a specific region of the mesoderm to the non-neural ectoderm and observed that an additional forelimb was formed. The original grafted cell population had the ability, even after transplantation, to “remember” its fate. Even if the cells were separated into two subpopulations and grafted independently, both grafts would grow an intact limb [19]. The key property of a morphogenetic field is thus to conserve its potential even after significant manipulation. Later observations showed that the fate of the morphogenetic field was dependent on the position along the antero-posterior axis. It led to the notion of *gradient field* which determined the identity of the morphogenetic fields [20].

Induction

In 1924, Hans Spemann and his student Hilde Mangold reported the discovery of a tissue in the newt gastrula that, when grafted on the ectodermal region of another newt embryo, triggered a neurulating process and initiated the formation of a secondary embryonic axis [21]. This tissue was called an *organizer*, as it was apparently able to instruct and organize the adjacent ectoderm. Spemann proposed two different speculative mechanisms: either the existence of a chemical substance that would be transmitted to the induced tissue, or the inducing tissue would possess a specific vitalistic “structure” associated to the living embryo [22]. These hypotheses became the focus of intensive study and debate.

Tissue boiling, desiccating and killing experiments practiced on the famous “Spemann organizer” rapidly dismissed the second hypothesis. A variant upon the theme of induction was introduced by Waddington in his work *Organisers and Genes*[23], in which an evocator-competence system defined an inducing substance, the “evocator”, that was only slightly perturbing the dynamics of the competent tissue, which would actively respond by a change of state controlled by the genes [24]. The organizer was no longer believed to actively organize the formation of the induced organs but only release a water-diffusible chemical agent initiating the self-organization of

the induced tissue. A global quest for the identity of the inducing substances started [25]. In 1932, Johannes Holtfreter used dead or desintegrated organizer tissue that still induced neurulation [26], and multiple chemical substances were diffused into the competent tissue [27]: lipids [28], oleic and nucleic acids [29], proteins [30]. As it became evident that induction was occurring in multiple tissues [31], ubiquitous candidate substances were targeted. In 1961, Lauri Saxon showed that the inducing substance could act through a “millipore” filter with an average pore size of 0.8 micron and a thickness of 20 microns, confirming that the substance was indeed diffusive [32].

Decades later, the great variety of candidate substances that positively induced neurulation progressively discouraged the embryologists to pursue their quest for tissue-inducing agents, while the more promising field of modern molecular biology was attracting the younger generations away from this problem [22]. However, it was only a temporary abandonment as the concept of induction witnessed a rebirth and was eventually reinstated on new physico-chemical grounds. The discovery of secreted proteins acting in evolutionary well-conserved signaling pathways, such as the mesoderm-inducing protein activin [33], launched again the quest for “the inducer” of the Spemann organizer.

1.1.4 Developmental Genetics

In the early 20th century, embryology and genetics were both part of the larger field of *heredity* and were tightly entangled. A distinction was initiated in 1926 by the work of Thomas Morgan who proposed that, to avoid confusion, embryology would study the expression of the hereditary traits, whereas genetics would deal only with the transmission of those traits ([34], Chapter II). From this time on, other biologists tried to reunite both fields, which led to the emerging field of *developmental genetics*.

The first publications at the foundation of this field were the work of Gluecksohn-Schoenheimer in the late 1930’s, who interpreted the defect in the induction of the mouse notochord as the consequence of a mutation of the Brachyury gene [35], [36]. The result would not only pioneer developmental genetics but also propose a new methodology for the study of embryology. Instead of perturbing experimentally the development of the embryo and observe the consequences on its phenotype, mutant phenotypes were to be observed first and genetic causes had to be inferred from them.

This methodological dichotomy was later merged with the experimental generation of mutants selected by the observation of their phenotypes. This was especially the case of chemically induced random mutations in *Drosophila* by Nüsslein-Volhard and Wieschaus in the 1980’s [37], which eventually earned them the Nobel prize, or more recently in zebrafish [38].

Waddington was also an important defender of the importance of genes in development. According to him, genes act as “controllers” of cellular fate. By comparing

the development of mutated *Drosophila*, he observed that a presumptive tissue (the “imaginal disc”) would transform into a leg or an antenna according to the mutation [24]. He illustrated his view by the concepts of *epigenetic landscape* and *canalization*, which he compared to grooves and bumps guiding a rolling “ball” of cell fate on a hilly terrain. Behind the scene, i.e. “under the hill”, genetic interactions reshape the folds, hence orchestrate embryonic development (Fig. 1.3).

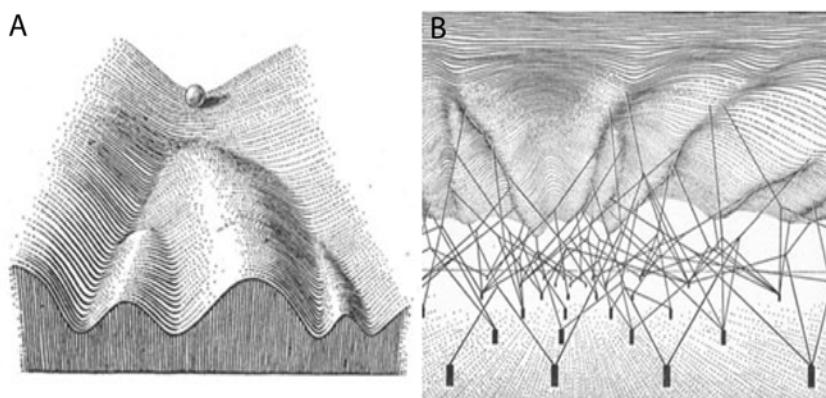


Figure 1.3: Waddington’s epigenetic landscape. A: The ball represents a cell evolving in the epigenetic landscape. Its fate is determined by the canals in which the ball is rolling. B: A view of the landscape’s behind the scene. The landscape relief is dynamically controlled by hidden wires that symbolize genes’ expression and interactions. Image and caption adapted from Slack [39].

1.1.5 Molecular Genetics

Operon-lactose

The discovery of the operon-lactose mechanism by Jacob and Monod in 1961 set the start of the genetic trend in embryology. It applied the idea of induction at the subcellular level by introducing genetic determinants, the *regulator and operator genes*, which explain how the rate of protein synthesis was controlled by the action of *repressors*[40]. This seminal paper already envisioned the influence of this discovery on embryology:

“The occurrence of inductive and repressive effects in tissues of higher organisms has been observed in many instances. . . It has repeatedly been pointed out that enzymatic adaptation, as studied in micro-organisms, offers a valuable model for the interpretation of biochemical co-ordination within tissues and between organs in higher organisms. The demonstration that adaptive effects in micro-organisms are primar-

ily negative (repressive), that they are controlled by functionally specialized genes and operate at the genetic level, would seem greatly to widen the possibilities of interpretation. The fundamental problem of chemical physiology and embryology is to understand why tissue cells do not all express, all the time, all the potentialities inherent in their genome.”

Immediately understood by some embryologists as Waddington, who had defended the notion of a cytoplasm-activated genetic control of cell fate in development (Chapter “The Activation of Genes by the Cytoplasm” in *Principles of Embryology*, 1956 [41]), this discovery opened the door to the reconciliation between the embryological orchestration of spatiotemporal cell specification and the biophysical molecular paradigm.

Gene regulatory networks (GRNs), *cis*-regulatory systems

The modern view of the orchestration of cell behavior in space and time is conceptualized by the work of Eric Davidson on the *cis*-regulatory system. It is an extension of the work of Jacob and Monod that systematizes the role of the genetic regulators as arrays of transcription factors’ target sites on the DNA [42]. These arrays, called the *cis-regulatory elements* as they are usually on the same DNA molecule as the genes that they regulate [43], define a network of interactions between the genes involved in development, called the *gene regulatory network* (GRN). The dynamics of this network is regulated by its topology and the quantities of the various *transcription factors* (TFs) that bind the *cis*-regulatory elements. As E. Davidson mentioned, the first GRNs were anecdotal but since 2002 rapid progress in their systematic analysis led to the publication of large-scale GRN maps, such as the network specifying the endomesoderm of the sea urchin embryo (containing over 40 genes), or the network responsible for the dorsoanterior-ventroposterior patterning and the endoderm formation of the zebrafish embryo [44].

Maternal factors

The dynamics of the GRN regulation is initialized by the various *maternal factors*[45]. The transcription factors are already present in the egg and serve as inputs to the GRN. The first direct evidence of a maternal RNA present in the oocyte and controlling the early activation of the GRN in the mouse was published in 1994 [46]. Maternal factor anisotropy is also an important cause of the patterning of the body plan. As the maternal factors are not homogeneously spatialized in the egg, a differential initialization of the GRN occurs in the different blastomeres. The *bicoid* gradient establishes the antero-posterior axis in the *Drosophila* embryo and is responsible for several asymmetries [47][48]. In the *C. elegans* nematode development, the *Skn1* maternal factor is concentrated on the posterior axis and specifies

the antero-posterior axis [49]. In the zebrafish, a maternal transcript, *Squint*, has been proposed as a predictor of the dorsal axis [50].

Signaling pathways, transduction

While the GRN dynamics view is centered on the cell, it also takes into account its communication capabilities through exchange of molecular information. Cell-cell interactions are realized by binding of secreted extracellular ligands, which trigger a transduction process and subsequent modification of the cytoplasmic dynamics. As mentioned by Pires da Silva and Sommer, a wide variety of cells use only a few classes of *signaling pathways*[51]. For example, among these pathways, the *Wnt* gene has been discovered multiple times in different animals. Its name itself is the contraction of two occurrences: the *Int1* gene characterized in 1982 by Nusse and Varmus that induced mammary gland tumors in mice [52]), and its homologue the *Wingless* gene (*Wg*) associated with *Drosophila* mutants that lacked wings [53]. The *Wg* mutation was later associated with defaults in the *Drosophila* segmentation process [37]). Other major signaling pathways are *Notch*[54][55][56], *Hedgehog*[57][58], *TGF β -BMP*[59], and *FGF* [60].

Mechanotransduction

Another important factor of cell dynamics is the integration of mechanical forces by the cell, or *mechanotransduction*[61][62][63][64]. Mechanical forces had been thought to play a major role in tissue and organ shaping since the end of the 19th century, from Wolff who studied the impact of the mechanical environment on the structure of bone tissue after fracture healing [65] to Roux [66] and D'Arcy Thompson [67]. More recently, forces have been demonstrated to influence vascular endothelial function [68]. The applied tension has been correlated with the proliferation rate in endothelial cells [69] and with the morphology of branching processes [70].

Evidence of mechanotransduction in modulated morphogenetic processes includes the role of forces on the cytoskeletal dynamics in the *Drosophila* mesoderm invagination [71][72] and the orientation of polarization axes in collective migration behavior [73]. An increasing number of studies stress the importance of mechanotransduction as an input of the GRN. The possible regulation of gene expression by microtubule-induced nuclear enveloppe fluctuations was shown in *Drosophila* [74]. A definitive demonstration of mechanotransduction's direct input into the GRN was given by Desprat et al. [75], who showed that compression forces exerted by a tissue could induce the expression of a transcription factor in another tissue. The corresponding TF, *Twist*, is involved in the differentiation of the anterior midgut in *Drosophila*. After removing the pushing cells by laser ablation, they were able to rescue *Twist* expression by experimentally applying forces with magnetic microtweezers.

Gene regulatory networks and epigenetics

In addition to mechanotransduction, a growing number of epigenetic features provide new types of inputs to the GRNs, introducing novel possibilities to control genetic expression during development. After the now classical DNA and histone methylation shown in different model organisms including the zebrafish [76][77], gene silencing by RNA interference is also described as a major concept part of the gene regulation machinery, also leading to a new kind of bioengineering tools [78].

1.1.6 Cell Biomechanics

The post-genomic era is bringing back the cell as the integrator of the molecular and genetic machinery. In this context, understanding precisely what the cell is doing is a major issue. Cell motility, cell adhesion, cell membrane deformation are all part of the biomechanics underlying morphogenetic processes and their emergent features at a macroscopic level. As reviewed by Ray Keller [79], this field remained quiet for a long period during the 20th century but was recently revived. Keller distinguishes between two tendencies that structure the physical shaping of embryos and were both envisioned by Johannes Holtfreter: the notion of “selective affinity” modulated by adhesion, and the notion of physical integration of multiple local cellular behaviors.

Differential adhesion hypothesis and improvements

Holtfreter employed his experimental skills to separate cells from their different germ layers and mix them. He observed that they were still able to recognize their lineage origins and adopt different preferential association or “affinities” accordingly [80]. He postulated that this mechanism could lead to the progressive organization of the embryo. In the 1960’s, Malcom Steinberg refined this idea and developed the *Differential Adhesion Hypothesis* (DAH) [81][82][83]. Comparing the behavior of cells during development to the properties of liquids, the DAH states that in a heterogeneous population cells are both cohesive and mutually motile, the interfacial surface tension leading the ensemble toward the most stable configuration. The main factors defining the interfacial surface tension were originally the mutual adhesiveness between cells, where higher affinity meant stronger bonds. This theory became very popular because of the simple causal link that it offered between gene expression and physical shaping through adhesion molecules. Later refinements added *cell rigidity* as a key factor to the interfacial surface tension definition. Through cortical tension, the driving principle became that stronger adhesion was increasing the contact size whereas stronger cortical tension decreased it [84][85][86][87].

Diversified cell behaviors

The second notion envisioned by Holtfreter was that cell mechanical behavior was truly diversified and that a global integration of the local behaviors had to be brought

to light. He observed the specialization of the external layer of frog gastrulae, in particular its organization as a planar sheet in which specific cell shape changes must reflect specific cell mechanical behaviors. The so-called *epithelial* cells are characterized by a strong polarization between the interior side, or *basal* side, and the exterior side, or *apical* side. They exhibit strong cohesion at their lateral interface and form surface layers [88]. During embryonic development, epithelial behaviors may be temporary as epithelial cell may leave the surface layer and migrate toward different regions. These cells are called *mesenchymal* and the transformation is the *epithelial-mesenchymal-transition*. Holtfreter also observed the protrusive activities of these cells in culture, and the way they exerted forces on the substrate and oriented their migration. Trinkaus determined the migrating behavior of cells in the avian neural crest, echinoderm mesenchyme and teleost fish epiboly [89]. The collective behavior of mesenchymal cells was reviewed in [90][91][92][93]. These concepts are the basis for more recent quantitative approaches of cell biomechanics in culture and *in vivo* [94].

1.2 Integrating Developmental Mechanics and Genetics: The MECAGEN Project

The ambition of the MECAGEN project, to which this dissertation contributes, is to construct an encompassing *model of the multiscale dynamics of the early stages of animal morphogenesis*. This theoretical endeavor must be controlled experimentally by a cohort of original quantitative *reconstructions* of the developmental processes, essentially taking the form of an *image processing workflow* on the one hand (see Chapter 7 of this thesis), and *agent-based model and simulation* on the other hand (Chapter 3-5). In this approach, *embryonic development is construed as a self-organized phenomenon arising from a multitude of individual cell behaviors, including their genetically and chemically regulated, and regulating, biomechanics*. It is a fundamental research project that relies on an important software programming effort required to (a) extract relevant measurements from real data (obtained by microscope imaging, Chapter 7), (b) recreate virtual data (by simulation, Chapters 3-5), and (c) validate the simulation against the measurements (Chapter 8).

The embryo as a complex system

Typical systemic properties of living organisms, such as “homeostasis” and “auto-poiesis”, can only be understood through a *complex systems* approach of their underlying biological processes. Complex systems refer to objects composed of a great number and diversity of small elements (e.g. organisms made of cells, cells made of molecules), which interact locally in a decentralized and self-organized way to give rise to a rich repertoire of non-trivial collective behaviors. This phenomenon is also called “emergence”, referring to the fact that the higher-scale properties qual-

tatively differ from the lower-scale elementary features, and cannot be deduced from them (if only because of a combinatorial explosion). The complex systems viewpoint requires new methodological and experimental strategies, in the case of development the use of animal models chosen for their properties of accessibility, transparency and phylogenetic position.””

MECAGEN wants to confront directly the level of the complexity of living processes, something that biology has so far partly evaded in its traditional attempts to address the “function” of genes one by one, or dissect subcellular processes in isolated cultured cells. This historical avoidance of complexity was undoubtedly necessary at first, and the reductionist approach has provided critical descriptions of the components, their local interactions, and their context. But now that the pieces of the puzzle have been (more or less) well identified, it is ample time to try and *integrate* them all together at the level of thousands of genes and millions of cells in order to see the big picture of the growing organisms. Moreover, the decomposition approach has limits that can be overcome only by viewing the elements in the broader context of their interactions.

Supporting platforms and methodologies

Our approach to developmental complexity consists of a reconstruction of multi-scale dynamics from measures based on *in vivo* observations at appropriate spatial and temporal scales. The MECAGEN project branched out of the integrative biology platforms (formerly European projects) *Embryomics* and *BioEmergences*, which have pioneered the design of methods and algorithms for measuring and reconstructing the dynamics of multicellular development observed by microscopy. The concept is that biologists produce and annotate time-lapse series of organism development, while mathematicians and computer scientists process these images to reconstruct (*BioEmergences*) and model (MECAGEN) collective cell dynamics. This effort resulted in sophisticated software platforms capable of handling large amounts of 4D imaging data through a workflow of segmentation and tracking algorithms. The concept of “multiscale dynamics reconstruction” does not come from biology, but from disciplinary fields that use formalization. Moving in this direction can be done only within an interdisciplinary context, where there is agreement about a representation of the system, the acquisition of relevant data and the processing of such data to extract quantitative measures. Then comes the question of theoretical models and what is expected from them. The relevance of such models depends on their predictive power, and this power is measurable only by going back to the experiments (see Section 1.3).

Summary of objectives and particular position of this dissertation

In this context, the objectives of MECAGEN are to implement:

- (A) the quantitative multiscale reconstruction of the morphodynamics of the zebrafish *Danio rerio*'s early embryogenesis, from the egg to the beginning of somitogenesis (15 hours of post-fertilization development at 28 degrees Celsius), supplementing the qualitative descriptions of biological processes with observations in 4D (3D+time) and measured data that allow the statistical analysis of individual and collective cellular behavior.
- (B) the modeling of the gene regulation processes, cellular dynamics and biomechanical constraints that govern morphogenesis via coupling between a discrete and a continuous formalization. Model and experiments are joined in a feedback loop, in which the model is optimised and falsified by experimental trials of “gain” and “loss” of function.

By its interdisciplinary nature at the interface between theoretical and experimental biology, including the modeling of molecular and cellular dynamics and their multiscale integration, MECAGEN relies on various methodological tools, mainly in computer science and algorithmic tools (image processing, multi-agent systems), mathematics (image processing, dynamic systems, statistics) and the modeling of large amounts of data (parametric estimation, optimisation).

The present dissertation will be focusing for the most part on the *biomechanical* side (“MECA”, Chapter 3) of the project as not only was it a huge endeavor in itself, but it was also a prerequisite to the understanding of gene regulation and chemical signaling (“GEN”, Chapter 4, only covering the basics) and the coupling of both sides (Chapter 5, only sketched out). Genes play a role only inasmuch as there is a physico-chemical phenomenon to steer and control. As James D. Murray puts it [95]:

“However one chooses to ignore mechanics, nevertheless, presiding over every embryonic twitch and jerk are Newton’s laws. And whatever role chemistry and genetics play in embryogenesis, they must finally submit their programs for Newtonian execution. Therefore, we have adopted the philosophy that, since morphogenesis is—at least proximally—a mechanical event, it is reasonable to start analyses of morphogenetic processes by examining the forces that produced them, and then, working backwards, add chemistry and genetics as needed” (p314).

1.3 Methodological Considerations

In this section, and before we set out on describing the model and experiments in the next chapters, we first wish to lay out our tools on the bench, i.e. the methodological principles and workflow that are at the foundation of our project. In short, we defend the notion that the complex nature of the processes involved in a phenomenon such as vertebrate development requires the use of “*augmented*” tools

and strategies, built upon the classical experimental scientific framework. While this is not a dissertation on the philosophy of science, we felt it was nonetheless crucial to clarify the framework of our modeling and simulation endeavors, which constitute an attempt at tightly integrating experimental observations with theoretical models in order to unravel the physical mechanisms of embryogenesis [96]. To this aim, we will make short but, we hope, important preliminary remarks about the generation of hypotheses and models by scientists and the tools that they design to help them in this process, then discuss the notion of validation and quality of a model.

Accordingly, this section is organized as follows: we first comment on the position of a scientist-modeler with respect to the external reality, i.e. her/his object of interest, but also the rest of the scientific community; then we examine the tools that can be used to “augment” the three fundamental scientific steps of perceiving, conceiving, and manipulating, which together form a loop; finally, we ask what it means for hypotheses to be ultimately “validated” by a fitness function, which is essentially a measure of the discrepancies between the simulated and the raw data.

1.3.1 The Scientist in the Observation-Hypothesis-Experiment Loop

The Individual

Experimental science stages an *individual* and her/his *environment*, or “reality”. Like any other explanation-seeking activity, experimental science is characterized by three fundamental processes in a cycle: perception of the environment, generation of new hypotheses, and experimentation on the environment to test these hypotheses (Fig. 1.4).

1. The loop can be entered by the individual perceiving and observing her/his environment.
2. The observations made by this individual are then matched with the knowledge that s/he holds. Most of the time, observations conform to this knowledge and no particular reaction is elicited. Otherwise, a significant difference between the observations and what was initially expected by the observer triggers a “curiosity” signal that challenges her/his existing set of hypotheses and leads her/him to reconsider some of them. The cognitive processes by which s/he creates new hypotheses (e.g. analogy, inference, induction, abduction, or deduction) are not discussed here. Ultimately, “satisfying” hypotheses are the ones that can establish causal relationships among the observations. They can identify certain observations (the effects) as the consequences of others (the causes). Hypotheses also have a predictive value as they allow to extrapolate the behavior of the system when the causal factors are modified.
3. The “experimental” qualifier attributed to many domains of biology or physics

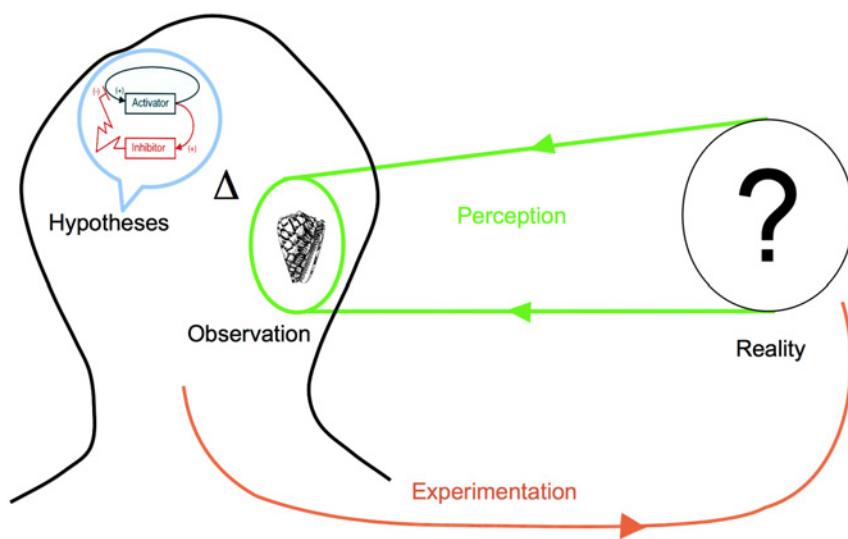


Figure 1.4: The observation-hypothesis-experiment loop of experimental science. It involves two main actors: the individual and reality (environment); and three main processes: perception, hypothesis generation (modeling), and experimentation.

comes from combining the pure thought exercise of generating new hypotheses and real interactions with external objects in order to test the validity of these new hypotheses. The most efficient way for an observer-modeler to assess that the causal relationships that s/he inferred are compatible with her/his observations is to identify some elements of the studied object as potential “factors”, then perturb these elements to modify the behavior of the studied object, and finally compare the new observed behavior with the predicted behavior. Note that the environment of the individual is made of multiple potential objects of study, so that the specification of one object of interest implies its separation from the environment, which may or may not include her/him. In developmental biology, the studied object is the embryo and the observer is excluded from the embryo’s natural environment.

Exchange/Validation by the Scientific Community

Even if the individual is at the center of experimental science, it is above all a *collective* knowledge-building enterprise. The interaction between an observer-modeler and the rest of the scientific community operates bidirectionally (Fig. 1.5):

- All the hypotheses made by an individual are elaborated upon a historical accumulation of prior scientific works. Today, s/he is potentially able to access all of the knowledge produced by the scientific community thanks to Internet,

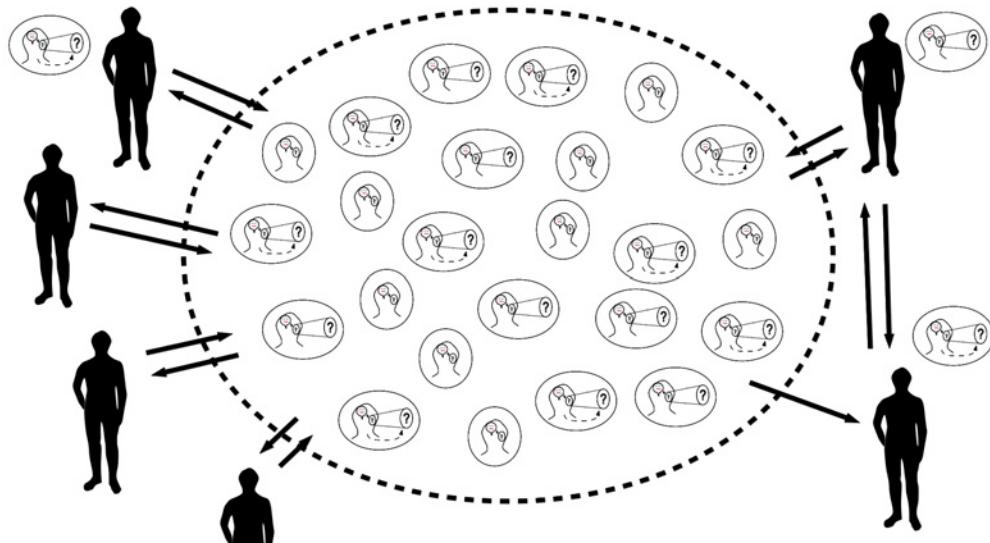


Figure 1.5: The collective effort of experimental science.

Each member of the scientific community may send or receive scientific work, which are verbal or written descriptions of parts or the whole observation-hypothesis-experiment loop.

in particular the various article databases (Pubmed, arXiv.org, IEEE, ACM, Google Scholar, etc.).

- One particularity of science is that the validation of a scientific work is ultimately decided by approval of the community. Through the peer-reviewed publication system, each new proposal is screened before being made available by a panel of individuals representing the community. This social dynamics is not without its problems, naturally (issues of motivation, expertise, time, politics, etc.), but consensus is basically the only mechanism that we have. We can distinguish between two types of peer validation: the *validation of the scientific work* containing all or some parts of the elements illustrated in Fig. 1.4, and the *validation of the hypotheses* contained in the scientific work itself. We will develop the latter aspect in Section 1.3.3.

1.3.2 Designing Tools to Perceive, Conceive and Manipulate

Experimental science insists on confronting hypotheses, the prediction they generate and observations. This confrontation is improved or “augmented” by the means of *tools*. Tools can be considered the third actor in experimental science, in addition to the individual and the object of study. In fact, they are objects of study in themselves. In developmental biology, the technologies used to observe (microscopy) or perturb (genetics, chemistry, mechanics) a growing organism are the focus of intensive research in other fields of science. The advances of our understanding are

closely coupled to the advances of these specialized and cutting-edge instruments. Microscopy imaging is constantly improving and expanding the spatiotemporal resolution and scope of observations. Every new microscope triggers a boom of new methodologies, observations and conceptualizations. For example, Fig. 1.6 illustrates the perception pathway augmented with such tools. We examine below three types of tools designed to augment the three fundamental processes of experimental science: tools to perceive, tools to conceive, and tools to manipulate.

Tools to Perceive

Tools can greatly improve the perception of the studied object, whether upstream at the level of the interface between the real system and the observed (raw) data, or downstream at the level of the “reconstruction” and analysis of this data to extract salient features compatible with its interpretation. Perception-augmenting tools allow to reach information inaccessible to the natural senses of the individual: they widen the scope of perception and increase its resolution at the same time.

Perception-augmenting tools, however, can also perturb the natural behavior of the studied object by introducing external elements. For example, in the present study, microscope images are obtained by using artificially mutated specimens of fish or injecting fluorescent molecules, which are also heated by the laser light that is designed to illuminate them. Therefore, special care must be taken to evaluate, control, and maintain the possibly deleterious perturbation to a minimum.

The goal of perception-augmenting tools is to produce *measures* of the studied object. Measures are a quantification of the physical attributes of the object by ordinary real numbers, the *data*, which are scaled in “units”. Therefore the data is the embodiment of the abstract notion of measure. Measures and data are never interpretation-free, as they are collected by choices that depend on prior knowledge. Thus their quantitative nature does not preclude subjectivity. They are indispensable resources that must be handled carefully and their scaling units validated by consensual agreement.

Optical Microscopy: An Interface with the Studied Object

In developmental biology, the principal perception-augmenting tools are obviously optical microscopes. The sets of measures used in the present project are originating from these devices. Other tools used in developmental biology, but more disruptive ones, are force microscopes and molecular biology techniques such as DNA/RNA microarrays, and macromolecule blotting and probing.

Optical microscopes extend our natural perception to the cellular scale and below. The general principle is to send photons to excite small regions of the embryos, which in turn emit other photons that are collected by camera sensors through objective lenses. The path of the exciting light beam can be controlled to cross the

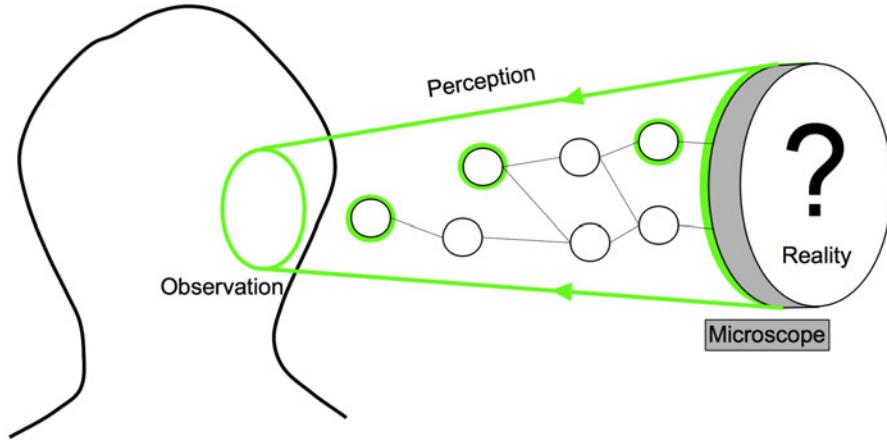


Figure 1.6: Augmented perception with interfacing tool (optical microscope) and a reconstruction workflow.

region of interest in the embryo. A software automates the task and automatically associates the spatial coordinates of the excited region to the quantity of photons captured by the sensors. An extensive scan produces a certain volume of “voxels” (3D pixels) that store the spatially localized quantity of photons emitted by the embryo. This process is repeated multiple times and a time-series of 3D volumes is generated, eventually producing 4D (or “3D+t”) *raw data*.

The value stored in a single voxel is called its intensity and belongs to a first category of raw data that we call here *local microscopic measures*, which are characterized by the smallest resolution both spatially and temporally. Then, the aggregation of all these local measures produces *extensive local measures*, a second category of raw data that can otherwise be called a “field” and contains the complete set of information captured by our perception-oriented device. Raw data presents two challenges:

- Its size is generally enormous. A few hours of embryonic development under the microscope typically produces billions of integer values. For example, 200 3D volumes of voxels of intensity sampled every 3 minutes, each volume having a resolution of $512 \times 512 \times 200$, yields over 10^{10} values. This size may also be multiplied by the number of channels used for light excitation. In the present study, two different channels are used for nucleus and cell membrane labeling. Multiple channels can be used to capture the light emitted by fluorophores that label gene expression [97].
- Consequently, profuse raw data is abstruse: it cannot be interpreted easily

and no biological insight can be gained from direct observation.

Phenomenological Reconstruction

Computers offer visualization software tools that allow the observer-modeler to create 3D+t movies of the captured developmental sequences. While a raw movie can certainly lead to qualitative insights, it also critically lacks quantitative measurements. Thus *data processing* is a necessary step toward a comparison with the predictions derived from the hypotheses, and a final interpretation. Here, we call this step the *phenomenological reconstruction*, or simply “reconstruction”, of the data. The idea is that, since raw data contains more or less incomplete information on the structure of the imaged embryos, depending on the spatiotemporal resolution and signal-to-noise ratio of the microscope, missing pieces have to be quite literally “reconstructed”. Moreover, as the reconstruction is realized from prior knowledge about the studied object, which is subjective with respect to the individual, this reconstruction is also “phenomenological”.

The reconstruction process is composed of a series of subprocesses organized in a *workflow*. Each subprocess carries out a specific task that extracts some information from the input data sets, completes it, and generates new data sets (Fig. 1.7). The question is then to define what type of measure the phenomenological reconstruction is aimed at. As described in Section 1.1, the individual cell’s dynamics is the fundamental unit of comprehension of biological development. The objective is thus to reconstruct the collective spatiotemporal dynamics of all cells taken together. The format we propose to use here is organized around the *lineage tree*, which follows along a global time line the complete cell genealogy starting from the zygote (Fig. 1.7). Each item of this graph represents a cell at a given time step. As we progress along the time axis, each item is connected downstream (a) either to a single item representing the same cell at the next time step, (b) or to two items if the cell has divided, where each item represents the daughter cells. The lineage tree is then enhanced or “decorated” by labeling each item with local observations about the cell it represents, describing its dynamics: its spatial 3D coordinates, membrane shape, list of neighbor cells, and various scalar quantities representing the fluorescent labeling molecules (RNA, proteins, etc.). In future work, we plan to add more precise information about asymmetrical quantities of labeled molecules, which represent cell polarity.

Equipped with this enhanced lineage tree, which we call the *reconstructed embryo*, the cell dynamics can be followed through space and time. The structured format of this exhaustive set of measurements allows comparison with predictions from the hypotheses, and among different reconstructed specimens. The reconstructed embryo is an “extensive local measure”, following the term we defined earlier. While it greatly facilitates and accelerates the handling of observations, the problem is that its dimensionality is also huge—as it scales with the number of time steps multiplied

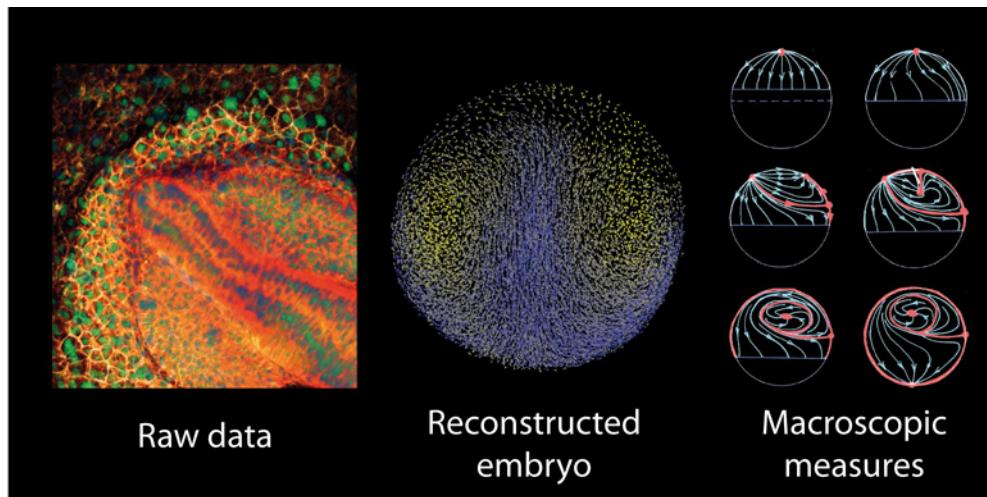


Figure 1.7: The three major steps of the reconstruction in the BioEmergences platform: raw data, reconstructed embryo and macroscopic measures. Left: raw data of the zebrafish developing head (nuclei labeled in green and cell membrane in red). Center: visualization of the reconstructed embryo with the MoveIt software. Small dots are the cell centers and arrows gives the future cell displacement. Colors indicates the velocities of the cell. The central image illustrates the extremely high dimensionality of the reconstructed embryo. Right: Macroscopic measures of the displacement field in the developing embryo (from B. Lombardot PhD manuscript [98]). Images realized by the BioEmergences team.

by the number of cells per time step multiplied by the number of local measures per cell. Thus it represents only the foundation of a higher-level reconstruction of the embryo dynamics, which ultimately provides the individual with interpretable biological facts. This higher level reconstruction represents a third category of data, which we call *macroscopic measures*. The macroscopic measures are obtained by projecting the enhanced lineage tree space onto specific low dimensional space characterizing biological properties. Performing multiple macroscopic measures allows a relevant characterization of the behavior of the global dynamics of the developing embryo.

Some of the reconstruction modules can be realized with off-the-shelf commercial software. However, the high number of cells involved and the difficulty to interpret and manipulate the 3D volume of data initiated in our group the design and implementation of better suited custom software and greater automation of the workflow of these subprocesses. Nadine Peyriéras at the NED (formerly DEPSN) in Gif-sur-Yvette and Paul Bourgine at the CREA lab and ISC-PIF institute in Paris, spearheaded two major European projects gathering several teams in six different countries: *Embryomics* (ended in 2009) and *Bioemergences* (renewed), which pioneered the design of such reconstruction methods and algorithms. While the biologists of the group produced and annotated time-lapse series of organism development, the mathematicians and computer scientists processed these images through specialized algorithms and transformations such as filtering, segmentation, detection, and tracking. This effort resulted in sophisticated software “*platforms*” capable of handling large amounts of 4D voxel movies of vertebrate embryos, and producing in output partial or complete cell lineage trees (see Chapter 7). For our part, we added new modules that we designed and implemented specifically for the present study. A detailed presentation of these workflows and our own contribution will be found in Chapter 7.

Tools to Conceive

Cognitive scientist Marvin Minsky provided a general definition of a model in his 1965 article *Matter, mind and models*[99]: “*To an observer B, an object A* is a model of an object A to the extent that B can use A* to answer questions that interest him about A*”. This definition is centered around the notion of “question” asked by the observer. The model is an object whose *raison d’être* is to satisfy its designer and, eventually, others around her/him. In this section, we discuss the nature of theoretical models, and particularly causal models. We believe that tools are also able to augment the capacity of the individual to make and test hypotheses, by providing interactions with a model that push her/him beyond her/his usual reasoning abilities.

Means of Expression

Models are constrained by their means of expression. The descriptive power of the structures and their interactions can vary greatly depending on whether they are expressed verbally, graphically or mathematically. Models are often described through a combination of the above. In the context of developmental biology, a distinction is often made between “classical” studies and “theoretical” studies: generally, the former use verbal and graphical formalism, while the latter use mathematical formalism.

Mathematical Formalism

In the mathematical formalism, objects are represented by variables, and their interactions are put into functions or equations. Generally, this formalism is expressed in “analytical form” using basic arithmetic operations such as $+, -, \times, \div$, power, exponential, logarithm, or infinite times series. In experimental science, the quantities involved tend to vary temporally and/or spatially, and their rates or derivatives play an important role. Typically, “ordinary differential equations” (ODEs) for time-varying quantities only or “partial differential equations” (PDEs) for time and space-varying quantities are used.

Dynamical Systems

In classical mechanics derived from the Newtonian laws, the time variable t holds a particular status. It is considered an “absolute”, meaning that two events are temporally separated by the same interval for all observers. This assertion is not correct in relativistic physics, where the notions of space and time are intermingled. However, the classical mechanics assumptions have founded a theoretical framework that produced accurate results for systems where objects are moving at a speed much smaller than the speed of light, or have sizes much larger than the atomic or sub-atomic scale (the realm of quantum mechanics)—which is obviously the case of all models of developmental biology discussed in the present work. Therefore, in the case of growing embryos that undergo spatiotemporal transformations, all theoretical representations fit well in the classical paradigm of *dynamical systems*. A dynamical system is built from three elements:

- the *state space*: the state of a modeled system being a collection of variable at a given time, the state space encompasses all the possible states that the system can adopt; it is defined a priori
- a set of equations that determine the laws of evolution of the system
- the initial state, i.e. the state of the system at the initial time from which the dynamical system evolves.

The interest of the dynamical system paradigm is that it does not restrict a model to a particular set of equations, but makes the space of possible values that variables can take and their initial conditions an important part of the hypotheses that define a model. Dynamical systems can be deterministic or stochastic if a random term (such as noise) is used in the differential equations. In a deterministic model, a state at any given time entirely determines a unique trajectory of future states of the system (Fig. 1.8).

Parameters

Certain variables have a special status as they are, by hypothesis, intended to remain constant along the state trajectories. In that case, they are called *parameters*. The design of a model must deal with parameters as much as with the mathematical laws of evolution. Selecting the parameters among the variables depends on the use of a model. Depending on the context, some parameters may be always fixed at a specific values because they have been confirmed by direct experimental measures. Parameters generally have their own space, the *parameter space*. Each point in the parameter space can be associated with a state trajectory (Fig. 1.8).

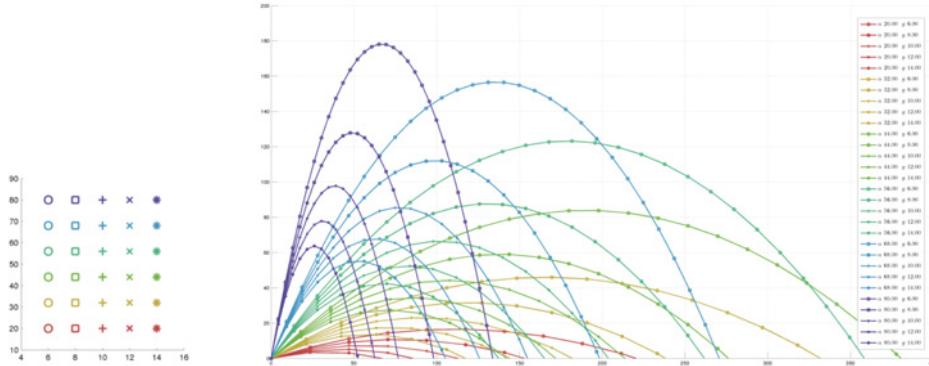


Figure 1.8: Simple illustration of a dynamical system: if a ball is hit by an object, it will move in space following a parabolic curve until it lands on the ground. A classical mechanics model would consider the curve (i.e. the temporal evolution of the position) as the “phenotype”, which is determined only from an initial known position and the velocity of the ball. Left: parameter space, Right: phenotypic space.

During the study of a model, a parameter may reveal itself as non-constant (we never ‘know’ if a parameter is truly constant, see validation section below). The reaction is to hypothesize a new rule for the evolution of the parameter and add it to the mathematical set of rule. This operation is a common part of the building of

model.

Theoretical Models, Analytical vs. Computer-Simulated Models

As mentioned above, theoretical models formalize the interactions among the system with equations that link together some of the selected variables of the studied phenomenon. Solving this *analytical* formalism is not always feasible because of constraints that are specific to mathematical symbolic transformations. Computers can help in this situation by converting equations into algorithms and calculating *numerical* solutions, which are approximations of the ideal solutions. Used nowadays in every field of research and engineering, numerical analysis allows scientists to tackle more complex phenomena. In 1952, Alan Turing already envisioned the use of computer to help him solve more realistic reaction-diffusion patterns in “The chemical basis of morphogenesis”:

“Most of an organism, most of the time, is developing from one pattern into another, rather than from homogeneity into a pattern. One would like to be able to follow this more general process mathematically also. The difficulties are, however, such that one cannot hope to have any very embracing theory of such processes, beyond the statement of the equations. It might be possible, however, to treat a few particular cases in detail with the aid of a digital computer. This method has the advantage that it is not so necessary to make simplifying assumptions as it is when doing a more theoretical type of analysis”[100].

Turing emphasizes the fact that the use of computer simulation is not only a practical solution to treat analytically unsolvable mathematical equations, but also that it allows the individual to integrate new mechanisms that s/he would refrain from using because of their mathematical unsolvability. In this sense, the computer (as a Turing machine) is a tool that augments the ability of the individual to develop mathematical models of the object of study.

An important category of analytically unsolvable models are called *many-body problems*, which concern most complex systems. They occur when a large number of elements are interacting together. As we will present in Chapter 3, the physical approach that we have chosen for our embryogenesis model is based on this assumption, each cell being an elementary particle interacting with its neighbors. Solving this system of equations is highly computationally intensive and requires the use of many computing units in parallel, such as computer clusters or graphical processing units (GPUs). In fact, computers were originally invented to deal with these situations (for example, the MonteCarlo simulation performed on the MANIAC computers in the early 1950’s [101]). Figure 1.9 illustrates the process of transformation of the model, from the original hypotheses made by the individual, to its theoretical form, and finally to its final form as a computer program.

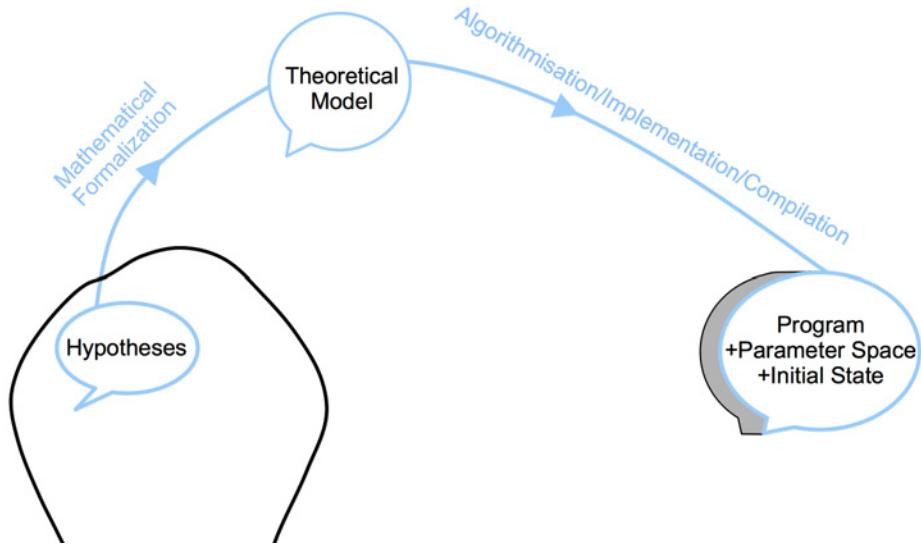


Figure 1.9: Augmentation of hypothesis generation through theoretical model and computer simulation.

“Black Box” Models

The type of models that have been considered so far are what some would call “white box” models: they are elaborated from specific hypotheses, and try to provide a mechanistic explanation. In contrast, “black box” models are empirical models built without *a priori* knowledge. These data-driven models are essentially *statistical* and used for data mining purposes. In our study, we have ruled out this category as it does not follow the framework of individual-induced hypotheses described above. Black-box models certainly possess predictive capabilities but they have little or no explanatory value, since the internal rules that they form are not directly interpretable. To this category belongs machine learning, such as neural networks or support vector machines, and evolutionary computation. An intermediary category, “gray box” models may be defined if partial knowledge of the system is included in the black box.

In conclusion of this section, as stated by Turing when introducing his reaction-diffusion model, a model is a “simplification and an idealization, and consequently a falsification”[100]. This description applies to all models, whether theoretical or “classical”, but mathematical formalism obviously still plays a fundamental role in at least three ways:

- Predictability: theoretical models allow to test hypotheses and postulate the behavior of complex systems in a way that “pure thought” could not. They are not a replacement for the experimental part, but they can perform theoretical pre-experiments to specify the conditions of “real” experiments.

- Abstraction: the idealization process allows to simplify the hypotheses and determine which mechanism is essential and which one is not.
- Precision: however sensitive qualitative transitions between different regimes or behaviors of a system may be, mathematical formalism can be adapted to any scale.

Tools to Manipulate

A third category of tools were designed to interfere with and modify the “natural” behavior of the objects of study, or their environment, in a controlled manner. Such experiments are artificial constructions that are designed to discriminate and select among various hypotheses about the rules of behavior of a system. In developmental biology, the embryo can be perturbed in two major ways: genetically or mechanically. Genetic experiments consist in making the embryo express an abnormal phenotype, either by random mutagenesis or by “morpholino” injection (antisense oligonucleotide morpholino modified to achieve specific gene knock-out by preventing RNA translation). Mechanical experiments can be done either through a lesion applied to a specific tissue to study its fate (e.g. laser ablation between individual cell-cell boundaries [72][102], and tissue dissection by laser [103]), or by some mechanical constraint to measure the response of the tissue. “Mechanotransduction” mechanisms also allow to conceive experiments at the interface between genetics and mechanics, such as provoking new genetic regulation by exerting a force with magnetic tweezers or magnetic nanoparticles [75]).

As it will be discussed in Section 2.2, studying an object often begins with studying its parts. In developmental biology, *in vitro* experiments allow to isolate cells or tissues, and test their behavior under controlled conditions (e.g. cell sorting experiments). One pitfall of this approach can be underestimating the impact of the artificial conditions on the behavior of the part, compared to its usual *in vivo* conditions. This can lead to the design of more complete *in vitro* environments that try to mimic and recreate the natural cellular “habitat”, as is the case for stem cells [104].

1.3.3 Reality Check: Validating the Hypotheses

The concept of *validation* of a hypothesis, which is employed in this work, can not be understood in the same sense as stating that an hypothesis is definitively true or false. Oreskes [105] has demonstrated that establishing the truth of a proposition is possible only in a closed system, and that models using incompletely known input parameters as is the case in developmental biology are never closed systems. Popper [106] also advocates that one cannot “prove” theories and laws, and that they can only be “falsified”. Thus in our case, “validation” can only mean a certain degree of *consistency* between the output of the model and the observations made about

the object of study. Observations can support the likelihood of a model [105], or its empirical adequacy [107].

The goal of an explanatory model is not merely to reproduce observations (as in the black-box methods) but rather unravel the principles that are at the foundation of these observations. The more observed data are positively confronted to the model, the more “adequate” the model and its underlying principles are deemed. The diversity of the observed data is another factor in favor of the adequacy of the model. A framework must also be designed to practice and confront the model against the observations of the studied phenomenon. The strategy that we adopt here is to *integrate the simulation platform and the reconstruction workflow*. In the same way that a microscope produces “real raw data”, our program generates “simulated raw data”. Then, the same reconstruction step is applied to both branches in parallel, giving rise to “reconstructed real data” and “reconstructed simulated data” (where reconstruction may refer to local microscopic, extensive microscopic or global measures; see above), which are later compared to each other. The reconstruction workflow gets a new leg from the theoretical process of our general experimental science scheme (Fig. 1.10). The different natures of the simulated and experimental data require that reconstruction algorithmic modules be applied so that they can be compared based upon the same format and automatically processed.

Fitness function

The comparison between simulated and observed values is based on a “distance” between dynamical trajectories, which can be embodied by a *fitness function* and applied to any one of the three categories of measures presented above (local, extensive or global). We distinguish between two types of fitness functions, in addition to the original “cognitive” comparison represented by the symbol Δ in Fig. 1.10:

- An *automated fitness function*, denoted by Δ_a : this function requires a reconstruction strategy based on the data generated by the simulation platform, similar to the reconstruction workflow described in the augmented perception part. It produces a quantitative score evaluating the discrepancy between two trajectories.
- A *visual fitness function*, denoted by Δ_v : the goal of this function is to support the individual’s intuitions and hypotheses based on visualization only. Visual fitness is not as formalized as automated fitness, but it constitutes an important stepping stone toward automation (a continuation of the perception-augmenting tools) and was extensively used in this project.

Validation Schemes

There are at least two different scenarios of exploitation of this integrated platform, as the design of the model and its comparison with the observations are tightly

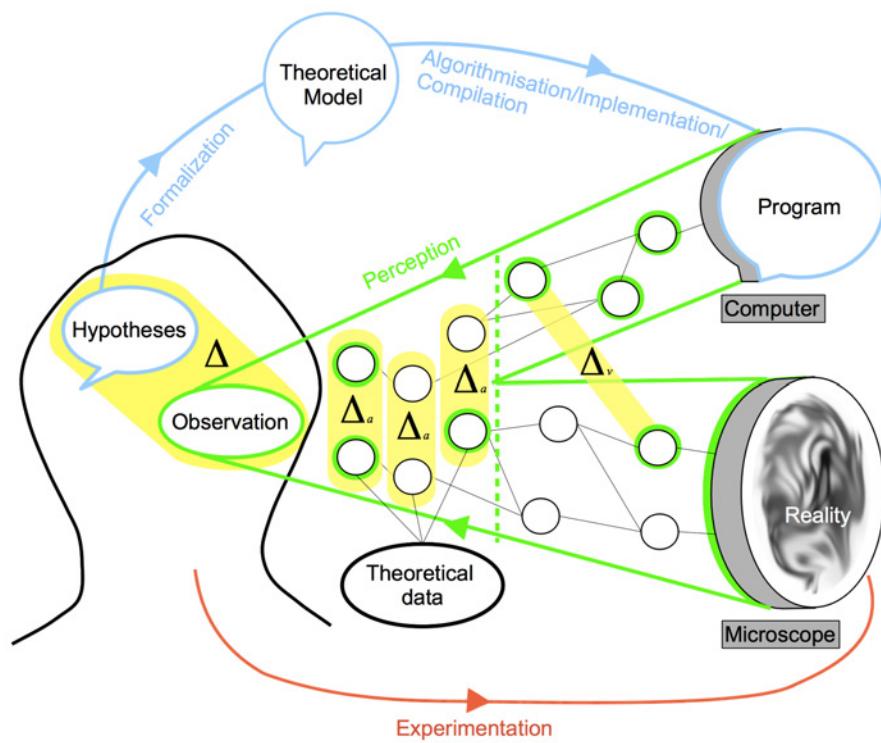


Figure 1.10: Integrative schematic of the tool-augmented observation-hypothesis-experiment loop.

coupled:

- asserting the likelihood of the model by showing that the foundational hypotheses of the model are sufficient to mimic the observations in a satisfying manner
- conversely, assuming the validity of the model and optimizing its parameters, then using its predictive abilities to design new experiments.

This scheme can be easily generalized to compare (Fig. 1.12):

- two or more models
- two or more individuals within a cohort (population of experimental individuals with the same a priori initial state, including genetic and environmental conditions)
- individuals from different cohorts
- models and theoretical data plotted in the macroscopic measure space.

An orthogonal distinction among fitness functions focuses on the type of observed data and simulated data that are compared (Fig. 1.13):

- Reconstruction of experimental raw data: we call this fitness function *experimental reconstruction fitness* (ERF).
- Theoretical data representing an idealized phenotypic behavior: we call this fitness function *theoretical fitness* (TF).

1.4 Overview of this Dissertation

The remainder of this study is organized as follows:

- **Chapter 2:** After the introduction chapter, which offered a summarized historical timeline of experiments and observations in developmental biology, we review next a few important families of formalized *models* of embryogenesis involving mathematical analysis and/or computational simulation. They include reaction-diffusion, morphogen gradients, epithelial cell shaping and cell sorting. Far from being exhaustive, this review only intends to be a sampler of particularly relevant papers that illustrate typical modeling paradigms. Then, we extract common principles from these various methodologies and attempt to unify them toward an encompassing modeling framework of multicellular development—which is the goal of the MECAGEN project.

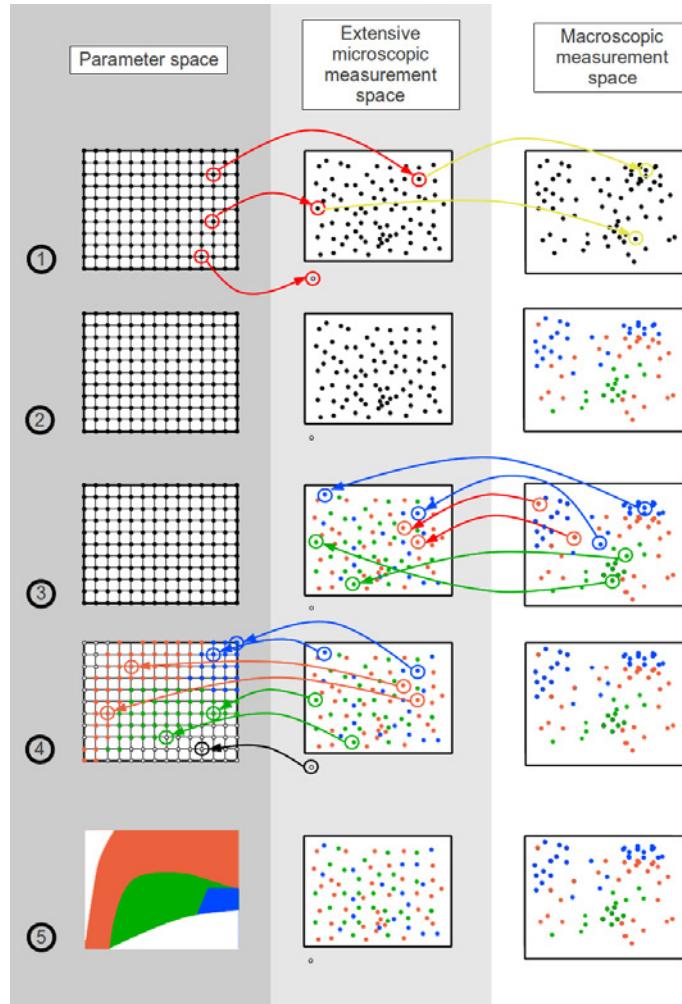


Figure 1.11: Generic evaluation of the performance of a model through quantitative evaluation of its parameter space. To each element of the parameter space (left column), a deterministic dynamical system model associates a trajectory that may be represented by a single point in the extensive microscopic measurement space (this is the state space but we want to emphasize the complexity and the high dimensionality of this space when the model is agent-based, center column). The macroscopic measurement point are obtained by reconstructing and measuring the observations of interest in the complex trajectory space. Row 1 illustrates the passage from the parameter space to the extensive microscopic measurement space (red arrows). As shown by the arrow pointing to the white dot, there is no guaranty that every point in the parameter space produces a “viable” trajectory. Quantitative assessment of the performance of the trajectory are determined according to adapted criteria (comparison with a target measure Fig. 1.12, clustering). A “color” is attributed to the macroscopic measure points to symbolize this evaluation (Row 2). The color is easily propagated back to the parameter space and allows a quantitative assessment of the performance of every trajectory (Row 3, 4, 5). White regions represent parts of the parameter space that are not viable (Row 5).

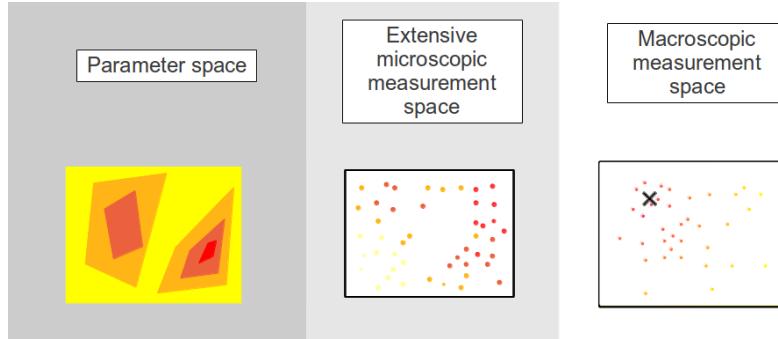


Figure 1.12: Fitness landscape according to a target macroscopic measure. This schema is a particular example illustrating the case of a model evaluated by the comparison a macroscopic measure (black cross). The score of each trajectory is attributed as a function of the distance between the simulated macroscopic measures and the target macroscopic measure.

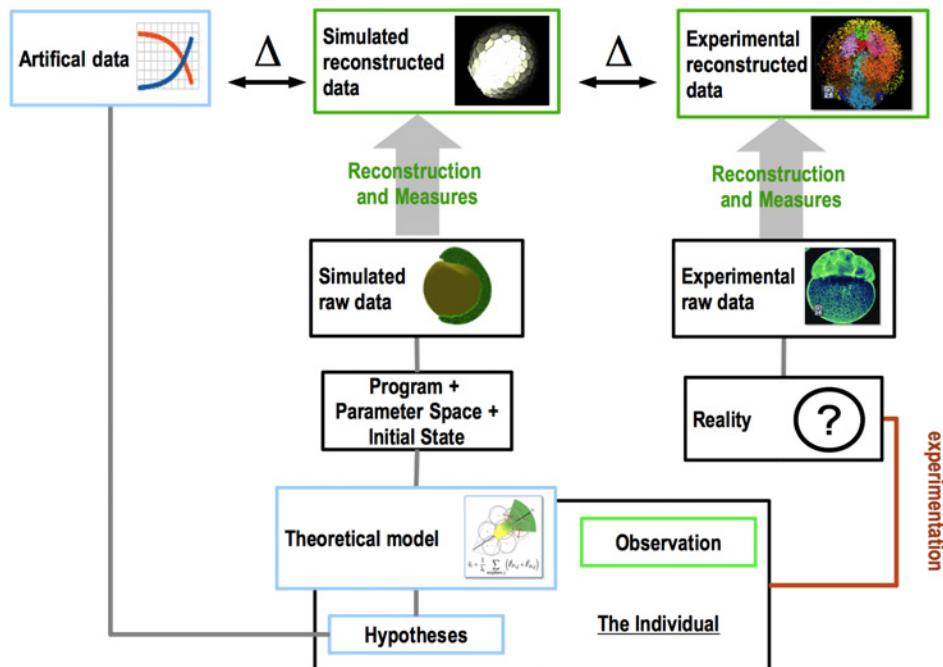


Figure 1.13: Summary of the methodological workflow adopted for this project. The top left Δ symbolizes the theoretical fitness (TF) and the top right Δ the experimental reconstruction fitness (ERF).

In the next three chapters, we explain the particular choices that we have made in the design of our model contributing to the objectives of the MECAGEN project:

- **Chapter 3 expresses and calculates the mechanical interactions and behavioral properties of the cells.** It presents a discrete-element model using one particle per cell driven by an overdamped equation of motion. Forces are calculated on each cell by summing over a mixed metric/topological neighborhood containing the nearby cells in contact with it. Two types of forces are involved: *relaxation* forces derived from an attractive/repulsive, elastic-like interaction potential, and a *behavioral* force moving the system of cells away from equilibrium.
- **Chapter 4 deals with chemical signaling and gene regulatory networks.** The goal is to briefly explain the principles of gene regulatory networks (GRNs), describe the components of GRN models, and give examples. In the present work, our particular objective was to design a simple and easily computable model of the molecular and genetic interactions that occur during development. Our model is articulated around three types of rules: rules driving the dynamics of intracellular gene/protein reactions, rules driving the dynamics of cellular secretion and transduction and rules driving the dynamics of extracellular reactions, transport and diffusion.
- **Chapter 5 lays out the first steps toward building a complete morphogenetic platform integrating mechanics and genetics,** as envisioned by the MECAGEN project. A simplified “cell behavior ontology” (CBO) is proposed. It relates *cell states*, determined by the concentrations of certain proteins, to *cell behaviors*, determined by biomechanical parameters.

While Chapters 2-5 above treated models of development from a generic point of view, Chapters 6-8 focus on the zebrafish early embryogenesis, including a short monograph of this species, a description of the raw imaging data reconstruction platform BioEmergences, and the results of our modeling and simulation work (“reconstruction of simulated data”) across several case studies:

- **Chapter 6 identifies the different phases of the developing early zebrafish embryo that our model should account for,** and the different components with their characteristic scales that are expected to be at play and underlie the biomechanics of this process. The zebrafish gastrulation, its processes and underlying causalities have long been and remain today a very active field of research. The global view that we have formed about this phenomenon, and which we present in this chapter, is based on raw microscope observations, the 3D+time imaging and reconstructions performed by the BioEmergences platform (explained in Chapter 7), and the state-of-the-art literature.

- Chapter 7 explains the software platform *BioEmergences* (created by our team, and extended by this study) that processes large amounts of 4D imaging data through a workflow of segmentation and tracking algorithms. This integrative biology endeavor has pioneered the design of methods and algorithms for measuring and reconstructing the dynamics of multicellular development observed by microscopy. Biologists produce and annotate time-lapse series of organism development, while mathematicians and computer scientists process these images to reconstruct (*BioEmergences*) and model (MECAGEN) collective cell dynamics.
- Chapter 8 showcases our modeling and simulation approach across six special studies of the zebra fish early development: (1) the yolk biomechanical properties, (2) the cell proliferation rate along the cell lineages, (3) the shaping of the blastula, (4) cell behaviors in the enveloping cell layer compartment, (5) intercalation patterns, and (6) gastrulation. In each case study, measures extracted from our “*in silico*” embryo are confronted to the image processing reconstruction provided by the *BioEmergences* platform. The goal is to tune parameters in order to validate our simulations and draw biologically meaningful conclusions.

Finally,

- Chapter 9 offers a discussion and conclusion

Part I

A GENERIC MODEL OF MULTICELLULAR DEVELOPMENT

CHAPTER 2

State of the Art and Principles of Embryo Development Models

In Chapter 1, we offered a summarized historical timeline of experiments and observations in developmental biology. Although the cited works made theoretical assumptions and proposed hypotheses about the mechanisms of development (Waddington, Needham and others had founded the “Theoretical Biology Club” in 1930), they did not contain formalized models *per se* with mathematical analysis or computational simulation. No symbolic or numerical calculation was performed. In the meantime and especially recently, an increasing number of theoretical and quantitative models of development have been proposed. We review here a few important families of approaches to embryogenesis such as reaction-diffusion, morphogen gradients, cell shaping or segmentation. Far from being exhaustive, our review only intends to be a sampler of particularly relevant papers that illustrate typical modeling paradigms. Then, we extract common principles from these various methodologies and attempt to unify them toward an encompassing modeling framework of multicellular development—which is the goal of the MECAGEN project.

2.1 A Review of Theoretical Models of Development

2.1.1 Early Attempts

Nicolas Rashevsky is one of the founders of theoretical biology. His whole research effort was spent on formalizing biology using mathematics. For this new discipline, he coined the name “Mathematical Biophysics” in 1938 [108]. He was the first to attempt a quantitative description of biological phenomena, and insisted on bringing theoreticians and experimentalists together to maintain close collaboration ties between them. This early interdisciplinary attitude already had its critics, as some said that he could produce “neither good mathematics nor good biology nor good physics” (from [109]). Concerning multicellular development, he proposed a framework to describe organisms as graphs in which vertices represented biological functions and oriented edges the interactions between them [110]. In particular, he showed as early as 1940 that cell polarity was possible even for cell with a spherical shape [111]. One of the earliest computer simulations of vertebrate development can be attributed to Jacobson and Gordon in 1976 [112]. They designed and calculated

a mathematical model of neural plate formation.

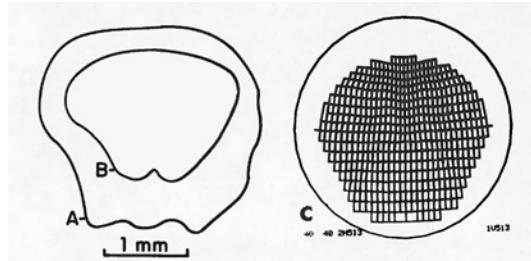


Figure 2.1: One of the earliest computer simulations of spatial transformation during vertebrate development.“A. Appearance of the neural plate immediately after isolation without notochord at stage 13. B. Shape attained in the same explant when controls had reached mid-neurula stage 15. C. Results of a computer simulation using shrinkage alone, which would be compared to B”. Images and caption from Jacobson and Gordon, 1976 [112].

2.1.2 Reaction-Diffusion Systems

Another historical landmark in the advent of developmental models was established by Alan Turing in 1952 with his breakthrough paper on “The chemical basis of morphogenesis”[100]. He proved that spontaneous order, such as stripes and spots of alternating color, could arise from the amplification of unstable fluctuations in an initially homogeneous substrate. The diffusion of two chemical signals or “*morphogens*” across a biological substrate, coupled with local chemical reactions is sufficient to generate complex spatial patterns. Turing predicted six stable states in his model depending on the reaction coefficients and the rates of diffusion of the morphogens. This idea was further developed and popularized by Gierer and Meinhardt in the 1970’s [113]. They showed that by combining “a short-range positive feedback with a long-range negative feedback”, they could generate all Turing patterns. Typically, a pigmented medium such as an animal coat undergoes spontaneous symmetry-breaking by diffusion of an activator substance and an inhibitor substance with two different characteristic distances [114]. This was also demonstrated by Young in 1984 [115] in an abstract model of vertebrate skin patterning implemented in cellular automata, in which the random initial distribution of pigmented and non-pigmented cells determined the final equilibrium pattern. An example of more biologically realistic reaction-diffusion model can be found in Amonlirdviman et al. [116], who propose that the *Drosophila* epithelium is polarized with the “planar cell polarity” (PCP) pathway in a non-autonomous manner, and in Lagendijk et al. [117], in the context of the cardiac valve formation. Reviews of reaction-diffusion systems can be found in Kondo et al. [118] and Murray [119][95].

Although reaction-diffusion models can theoretically account for all sorts of patterns, the biological stripe patterns where this framework indeed applies are much more rare. It appears now that the stripes gene expression patterns in Drosophila segmentation cannot be explained by reaction-diffusion models [120]. The zebrafish pigmentation offers another contrasting example of pattern that does not form through reaction diffusion *stricto sensu* i.e. with a single molecule diffusing at long range, but rather via a “combination of other signaling mechanisms that have long and short functional distances”[121]. More generally, the reaction-diffusion models rely on long-range diffusion and at first are not made to account for the cellular organization of the tissues, their relevance in modeling morphogenesis and patterning in multicellular organisms is much more limited than anticipated by theoretical biology.

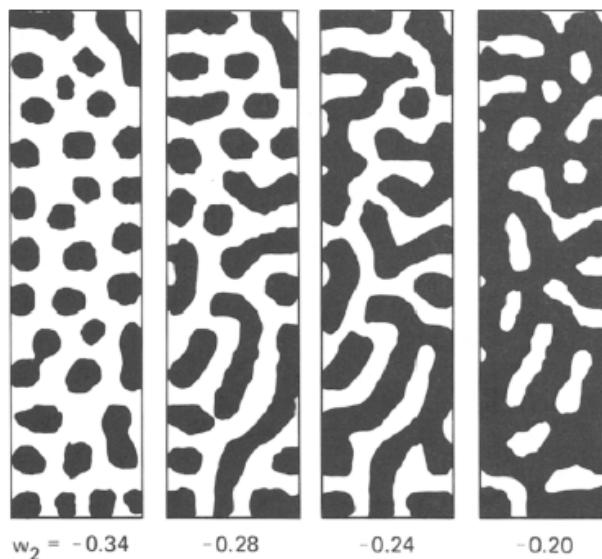


Figure 2.2: Patterns produced with the activator-inhibitor model. As inhibition is decreased (left to right), the spot pattern connects up into a pattern of stripes. Images and caption from Young, 1984 [115].

2.1.3 Morphogen Gradients and Positional Information

In reaction-diffusion models, the key concept is that elements (at the molecular or cellular level) interact with their neighbors purely locally to produce ordered but repetitive patterns of contrasting activity (e.g. chemical species or pigmented cells). This family of phenomena is commonly referred to as *pattern formation*. Complex patterns produced by nature have always been a source of great fascination to philosophers and scientists: ripples in sand dunes, spots in animal coats, geomet-

ric figures in plants, and a multitude of meanders, spirals, branches, and lattices observed everywhere. Most biological systems, however, distinguish themselves by strong morphological properties, i.e. an elaborate shape and body plan *architecture*, which are much more sophisticated than texture-like pattern formation. Striped and spotted patterns typically result from the amplification of unstable fluctuations. Setting aside questions about the actual existence of activator and inhibitor “morphogen” agents, it remains that the pattern formation phenomena covered by such models are fundamentally random and unpredictable. Are there going to be four, five or six spots? Although the patterns are often statistically homogeneous and can be described by a characteristic scale or order parameter (diameter of the spots, width of the stripes), morphological details such as position, orientation and number are not invariants of the system.

Thus Turing-like reaction-diffusion principles might be able to account for some pattern formation effects in biological development, such as mammal coat, butterfly wing spots, angelfish stripes [115] or seashell motifs [122], yet these effects seem only secondary or literally “superficial” compared to the overall form of an organism. The precisely arranged body shape of animals, made of articulated segments and subparts, is not the result of free-forming random instabilities. It is a fundamentally “guided” morphogenesis process that plays out under deterministic control from the genome. Except for very rare cases of malformation, all members of a pentadactyl mammal species reliably develop five digits, not sometimes four or sometimes six. All healthy embryos of *Drosophila* exhibit exactly seven bands of differentiated gene expression along the anteroposterior axis, which then give rise to 14 segments [37]. Each one of these mammal digits or insect segments is independently controlled by a specific combination of genes. At every time step in the development of an embryo, a homogeneous region of the overall embryo pattern is defined as a local group of cells that have the same gene expression profile, i.e., the same dynamic regime of RNA and protein concentrations.

In sum, biological forms are not statistically uniform. They are rich in morphological information and cannot be reduced to one characteristic scale like reaction-diffusion patterns. Some free pattern motifs (spots, stripes) can be embedded in a guided form (leopard, angelfish). Conversely, a guided form can be duplicated and distributed in free patterns (e.g. hundreds of copies of the same flower shape on the branches of one tree). Biological forms can thus combine a little free patterning with a lot of guided morphogenesis. The latter kind can be more effectively modeled through the paradigm of *positional information* (PI) introduced by Lewis Wolpert in the 1960’s [123] (later revisited and extended [124][125][126]) to account for heterogeneous and large-scale spatial patterns of cellular differentiation. At an abstract level, the key idea is simply that cells must establish long-range communication system that allows them to create to generate different parts of the organism in different locations. It is inevitable that some form of PI should be at work in multicellular organism development, embodied in various possible ways, either (1) through pas-

sive diffusion of morphogens spreading throughout the tissue and acting directly on distance cells and/or (2) through cell-to-cell intermediate-messenger signaling that would relay the signal [127][128][129] (Fig. 2.3). Several experimental techniques have been developed to study the modes of propagation of the morphogens [130]. Both mechanisms give rise to concentration *gradients*, either in the extra-cellular matrix (ECM) or in the cells' cytoplasm. In effect, PI is a genuine *coordinate system* that self-organizes by decentralized chemical signaling among cells. Recurring at multiple levels of details in a (non-self-similar) fractal fashion, it constitutes the basis of an entire “hidden geography” for the embryo, following Enrico Coen’s words [1], and is also employed in abstract models of development and artificial systems [131][132][133].

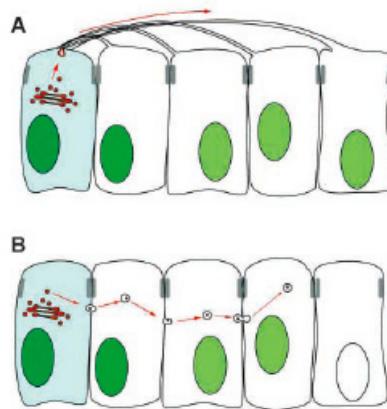


Figure 2.3: Models for morphogen transport. (A) A model for cytonemes. Cells at the periphery of the imaginal disc extend long processes, cytonemes, towards the AP border, where Dpp is expressed (light blue). (B) A model for argosomes. The basolateral membranes of imaginal disc cells vesiculate and travel throughout the disc epithelium. Image and caption from Tabata and Takei (2004) [129]

A famous abstract illustration of gradient-based morphogenesis is the so-called “French flag problem”, also coined by Wolpert [123], in which three domains of different colors (blue, white, red) representing three modes of cell differentiation arise on top of a decreasing morphogen concentration profile. When the morphogen concentration crosses certain thresholds, cells start expressing other genes, i.e. change types. This type of “programmed patterning” has been the topic of many models, whether on a fixed lattice background or in combination with cell division and motion [134][135].

However, a major issue with the reliance on concentration gradients for morphological patterning is obviously their *robustness*[136]. If they are supposed to

determine, control, or even only influence the layout of the body plan and the structure of the organs and appendages, then precision should be a critical feature. Yet, this precision will not be found in the gradient diffusion or signal propagation themselves, as they are a highly noisy and approximative process. Rather, robustness is an emerging property of PI *combined with* gene regulation (possibly via a correcting “attractor dynamics”[137][138]) within the continually changing spatial environment of the growing organism [139]. This puzzling question of the gradient-based patterns’ “scaling invariance” as the tissue changes its size has triggered a number of research works [140], in particular in the *Drosophila* early striping and segmentation. Various mechanisms have been considered, whether through the ratio of two opposite morphogen gradients [141] (Fig. 2.4), the degradation of morphogens by nuclei slowing down as the embryo expands [142], the ubiquitous production of a size-dependent regulatory element [143], or an “expansion-repression” coupling between the morphogen and an additional molecule [140].

As for the reaction-diffusion model, it should be noted that the ideal theoretical case of a morphogen gradient generated by diffusion is unlikely in a multicellular space. Rather than generating a gradient by diffusion, one has probably to account for cellular fields with graded intracellular expression of potential morphogens.

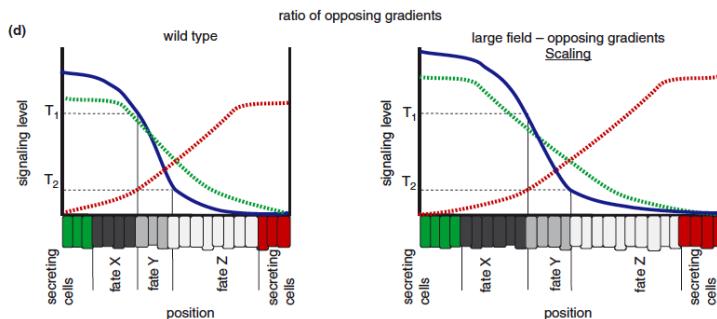


Figure 2.4: Scaling of morphogen gradients. (d) The ratio (blue) between two gradients emanating from opposing edges of the field (green and red) can provide scaling of a signaling gradient. Left panel: wild type field; right panel: larger field. Image and caption excerpt from [140]

2.1.4 Epithelial Cell Shaping and Division Patterns

A relatively recent interest in epithelial tissue modeling has generated a certain number of models focused on cell shaping, distribution of neighborhood sizes and division axis fields. In [144], the authors use a Markov chain model to explain the evolution of the distribution of cell shape in the *Drosophila* epithelium (Fig. 2.5A,B). They propose that cell proliferation, and not cell packing, is responsible for the shaping of cells in monolayered epithelia. The model is generalized and

compared to other organisms. Other investigators [145] contend that physical forces, in addition to cell division, are also required to explain epithelial cell shape in the wing disc of *Drosophila*. They use a vertex-based model in which vertices represent junctions between wing cells (Fig. 2.5C). Forces are derived from an energy function that takes into account cell elasticity, cortical tension and intercellular adhesion. The model is tested experimental data obtained by laser ablation. Patel et al. [146] investigate two factors potentially responsible for cell proliferation: inheriting the cleavage plane orientation from mother to daughter cells, and symmetry of the division. They conclude that strong symmetry is the dominant factor explaining the distribution of shapes observed experimentally. Other vertex-based models of cell junctions can describe plant growth. For example, the particularity of the shoot apical meristem of *Arabidopsis* is that it is stretched by an isotropic mechanical tension controlling the growth rate and division of the cell. A study of division patterns in this tissue [147] shows a result similar to the previous work: symmetric division is the favored factor explaining the observed cell shape distribution.

Sandersius et al. [148] investigate epithelium patterning before and during the primitive streak formation in the chick embryo. For them, against Gibson et al. [144], non-spatial Markov models are not sufficient to explain the histogram of number of neighbors in proliferating only epithelium. They argue that any attempt to improve biological plausibility of this type of model (e.g. with 3-sided cells or asynchronous division) induces a deviation from, instead of a refinement of, the “standard” histogram observed in various species. On the other hand, they show that their own geometrical epithelium model (based on a “Subcellular Element Model”; see Section 3.1) predicts the histogram with growth rate being the unique meaningful parameter. Escudero et al. [149] introduce complex network topological measures in addition to the cells’ neighborhoods and geometry. These new measures allow them to discriminate among epithelia belonging to different species, different stages of development, or genetic variants of the same species. The observed data is classified with statistical methods, allowing them to reveal the “signature” of an epithelium. The question of the interplay between cell shape and cleavage-plane orientation is reconsidered in [150].

2.1.5 Differential Adhesion and Cell Sorting

The concept of patterning and compartment formation through cell sorting by differential adhesion, which was developed by Steinberg under the *Differential Adhesion Hypothesis*, is both powerful from a theoretical point of view and for its adequacy to describe the biological systems in agreement with typical biological descriptions at the tissular, cellular and molecular level. The dissociation and gathering of mixed embryonic cells populations leads to systems quite similar to mixed immiscible fluids. Different phases tend to cluster together to eventually form two clusters, one often engulfing the other. Similarly to fluid behavior, Steinberg hypothesized that

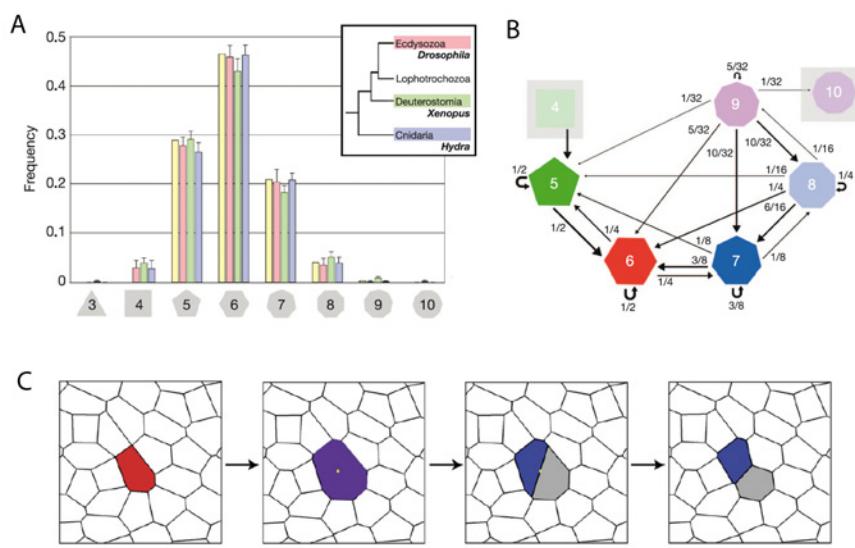


Figure 2.5: Selected illustrations from the literature of epithelial cell shaping and division patterns. A: Comparison of the distributions of polygonal epithelial cells obtained from simulation (yellow), the *Drosophila* wing disc (pink), the *Xenopus* tail epidermis (green) and the *Hydra* epidermis (blue) [144]. B: Schematized Markov chain model expressing the proliferation dynamics of polygonal cells and used in the simulation of A [144]. C: Typical cell division event in a vertex-based model [145].

cells should also minimize their surface area and that this process is modulated by differences in cell-cell adhesion [81][151][152][82][83].

Theoretical modeling and computer simulation in 2D and 3D of differential adhesion-based cell sorting have been repeatedly carried out. In 1992, Graner and Glazier published their seminal cellular Potts model, initially in 2D [153][154]. Other theoretical frameworks were used to simulate similar processes: Broadland and Chen favored a finite element model [155], Landsberg et al. developed a vertex model for a network of adherens junctions to simulate the formation of compartments in *Drosophila* embryogenesis [102][156], Beatrici and Brunnet explored the possibility to achieve cell sorting solely by motility differences in a model of self-propelled particles [157].

The relevance of the differential cell adhesion model has been reinforced by the characterization of cell adhesion molecules and their quantitative contribution to the surface tension in aggregates, thus providing a molecular basis to cell sorting [158]. The nature of the dependency between the surface tension of an aggregate and the cadherin expression level has also been explored theoretically [159]. More recent studies have tried to experimentally distinguish the respective contribution of cell adhesion and cortex tension in cell-cell contact formation, cell sorting and tissue segregation. Heisenberg, Paluch and colleagues approached these issues in zebrafish early development and the segregation of embryonic progenitors [87].

2.2 Common Modeling Principles: Toward an Integrated Theory of Development

Most of the models reviewed in the previous section are focused on specific aspects of development, whether certain episodes of embryogenesis localized in space and time, or particular mechanical or genetic components of the dynamics. The ambition of the MECAGEN project, and generally the modeling community, is to integrate all these dimensions into one comprehensive, or “*in toto*”, model. The benefits of stitching together various approaches and components are to push them toward a mutual adaptation and cross-validation in the perspective of building a global homogeneous and consistent framework. Typically, studies of local phenomena follow a “figure/background” template in which the object of interest is the figure and is modeled in detail, while the environment is a background and is only a rough approximation. In an integrated model, both are “figures” and are represented by rules and laws.

The ultimate goal is to unify what Darwin called the “endless forms most beautiful” of nature, and construe them as variants around a common theme [160]. The variants are the unique genetic and epigenetic information of each species; the common theme is the developmental dynamics that this information guides and “parametrizes”. The Modern Synthesis of evolution and genetics postulates this re-

duction in principle but has never truly modeled and explained it physically. While the attention was focused on selection, it is only during the past decade that analyzing and understanding *variation* (as the generation of phenotypic innovation) by comparing the developmental processes of different species, at both the embryonic and the genomic levels, became again a major concern of biology. This is the topic of *evolutionary development*, a recent and rapidly expanding field of biology nicknamed “evo-devo”. Researchers realized that the genotype-phenotype pairing could not forever remain an abstraction if they wanted to understand the unique power of evolution to produce countless innovative structures. To quote Kirschner and Gerhart [161]:

“When Charles Darwin proposed his theory of evolution by variation and selection, explaining selection was his great achievement. He could not explain variation. That was Darwin’s dilemma. . . . To understand novelty in evolution, we need to understand organisms down to their individual building blocks, down to their deepest components, for these are what undergo change.” (page ix)

Evo-devo casts a new light on the question still little addressed by today’s predominant gene-centric view of biology: To what extent are organisms also the product of complex physicochemical developmental processes not necessarily or always controlled by complex underlying genetics? Before and during the advent of genetics, the study of developmental structures had been pioneered by the “structuralist” school of theoretical biology, which can be traced back to Goethe, D’Arcy Thompson, and Waddington. Later, it was most actively pursued and defended by Kauffman [162] and Goodwin [163] under the banner of “self-organization”, argued to be an even greater force than natural selection in the production of viable diversity.

The grand challenge of creating a universal generic model of multicellular development, however, requires solving difficult problems: dealing with multiple levels of organization (metabolites, macromolecules, cells, tissues, organs), relating one level to the next by confronting the issue of “emergence”, and eventually identifying custom laws at each level—as the reductionist dream of a huge atom-based simulation is not conceivable or at best completely unpractical. It involves a mix of continuous and discrete approaches, combining analytical, statistical, and agent-based computational models. Several works have ventured to propose such integrated models at various degrees of completion and with different emphasis, whether a multiscale integration of mechanical forces [164][165], multiscale modeling of pattern formation [166][167], a multi-model framework for the simulation of morphogenesis such as the CompuCell platform [168][169][170], or multiscale abstract models and artificial life systems [171][172][173][174].

2.2.1 Dealing with Multiple Levels of Organization

Although “materialism” and “physicalism” have different historical roots, their meaning today is basically equivalent. They state that everything in the natural world is physical and conforms to the action of physical matter. Moreover, the study of biological development is well within the framework of classical physics and is not concerned with more speculative domains at extreme scales such as quantum physics (dealing with objects smaller than atoms), general relativity (dealing with objects near the speed of light), or consciousness (dealing with the phenomenology of the subject). Modern science’s understanding of the material world is traditionally categorized into a hierarchy of levels (Fig. 2.6). At the smallest scales reside subatomic particles and possibly superstrings, which create the *atoms*. The atoms arrange themselves into *molecules*, which represent the lowest relevant level for the study of biological development. Small molecules in turn interact to form macromolecules (DNA, RNA, proteins) and organelles. These objects constitute the basic structural units of life, the *cells*, which self-assemble into tissues, composing the organs and the *organisms*. Ultimately, through their social interactions, organisms form populations, ecosystems, and finally the biosphere, which represents the whole pyramid of life on Earth.

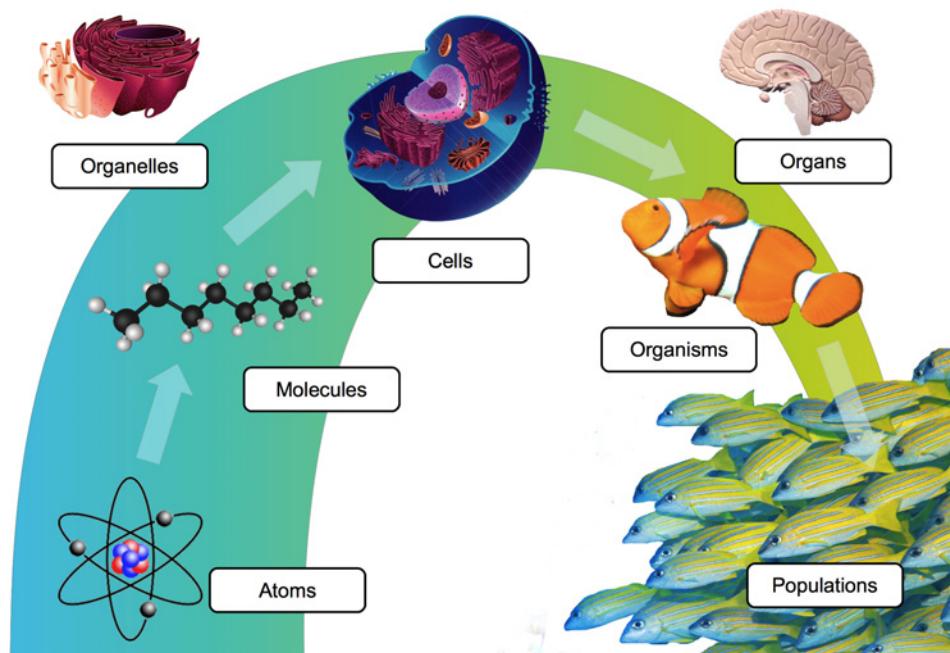


Figure 2.6: Biology is studied at many levels of organization.

At each level of organization, entities are *complex systems* composed of a great number of small, repeated elements that interact locally and produce a collective behavior in a decentralized and self-organized fashion. They share similar structural

and functional properties and are themselves internally structured as ensembles of smaller entities at a finer scale (Fig. 2.7). For example, one cell can be modeled as a self-regulatory network of genetic switches, one social agent (insect) as a network of decision rules. Conversely, agents also interact collectively at the level of clusters or subnetworks (organs, assemblies, cliques) to combine in a modular fashion and form larger sets. Thus, from both perspectives, complex systems can often be described as “networks of networks” on several hierarchical levels. The higher levels connecting elements or clusters of elements are generally spatially extended (cell tissues, ant colonies), whereas the lower levels inside elements are generally nonspatial (gene nets, rule sets). Elements follow the dynamics dictated by their inner networks and also influence neighboring elements through the emission and reception of signals (chemical). Practically, however, complex systems models rarely involve more than two levels of organization.

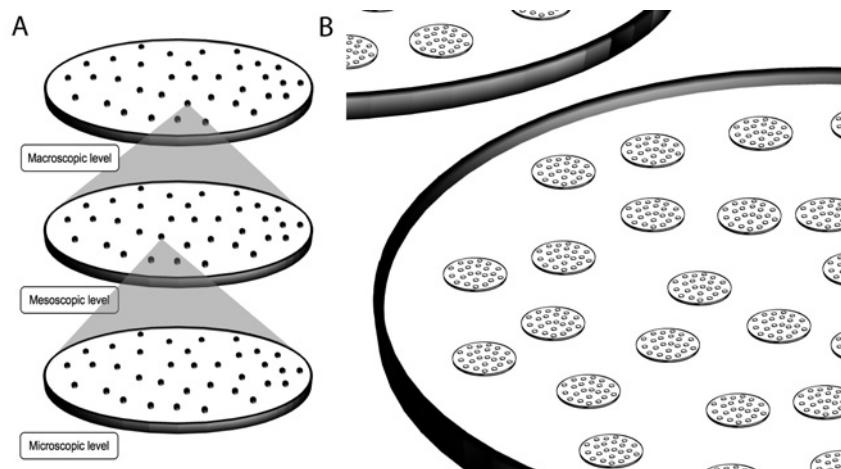


Figure 2.7: The hierarchy of level of organization of physical matter. At each level of organization, entities are *complex systems* composed of a great number of small, repeated elements. A: Level view. B: Integrated view.

Physicist Richard Feynman once declared that “If we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that all things are made of atoms, and that everything that living things do can be understood in terms of the jigglings and wigglings of atoms”[175]. Feynman epitomizes the *reductionist* school of thought according to which the behavior of any living thing can be explained exclusively in terms of its parts, subparts, and their interactions. This standpoint advocates a pure *bottom-up* view of causality. On the other end of the spectrum, following the famous words of Aristotle “the whole is *more* than the sum of its parts”, *non-reductionists* or *holists* claim that understanding the world, in particular life, must rely on conceptual structures at intermediate levels adapted to the studied phenomenon. This attitude is exemplified

by P.W. Anderson in this article “More is different”[176]. The central concept here is that of “emergence”, which as long as it has not been or cannot be reduced, justifies the very existence of all the scientific disciplines, from biology to social sciences, that are not just temporary placeholders for a yet-to-be “physics of everything”, but relevant in themselves.

Although our personal inclination is to agree with the reductionist view, i.e. living matter could be ideally understood in terms of its atomic components, the immensity of the number of atoms exceeds the capacity of the human brain to comprehend how this type of system works. We may rely on the ever-increasing computational power of computers, yet it might only be able to “reproduce” but not “explain”. Rephrasing Feynman, one could say that “everything that living things do could be ‘recreated’ in terms of the jigglings and wiggles of ‘simulated’ atoms”. A whole new field of research, generally known as “computational chemistry” or “particle-based molecular dynamics”, is attempting just that. Although still far from comprehensive realistic simulations, it has considerably progressed over the past few decades, boosted by the explosion of computational power. For example, Jensen et al. [177] simulate the mechanisms of voltage gating in potassium channels by a molecular dynamical model whose typical simulation can process 200,000 atoms in 10 million time steps, corresponding to a total simulated time of 200 microseconds (i.e. 2 simulated femtoseconds per time step). Latest hardware technologies based on general-purpose graphics processing units (GPUs) can support a molecular dynamics library such as the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [178], providing an estimated computational efficiency of 20 nanoseconds of GPU time per atom per time step (as of 2012). In this context, the above simulation of voltage gating would run about 100 hours on a single GPU, thus by extrapolation a single cell corresponding to simulated time of 10 hours would take 10^{25} seconds, i.e. a billion times the age of the Earth. Optimistic reductionists still rely on a continual acceleration of computing speed (possibly via novel technologies such as quantum computing) to considerably reduce this enormous duration and make it a practical tool.

2.2.2 Relating One Level to the Next: The Problem of Emergence

approche statistique: macro as an “averaged” approximation of the microscopic behaviors approche multi-agent: direct interaction of the microscopic “agent” -> macroscopic reconstruction ?

The dynamics of some complex systems can be described and formalized by macroscopic equations. Typically, the goal here is to solve these equations to obtain an explicit expression of the behavior of the system over time: for example, ordinary differential equations (ODEs) of macro-variables in well-mixed systems (e.g. the law of mass action in chemical kinetics), or partial differential equations (PDEs) of local variables in spatial extended systems (e.g. the heat equation or wave equation).

In certain cases, it is even possible to find the formulation of an exact solution by calculus, i.e. the symbolic manipulation of expressions (e.g. the solution of a geometric growth is an exponential function; the solution of the heat equation in homogeneous space is a Gaussian). Unfortunately, although vast, this family of analytically expressible *and* solvable systems is in fact very small compared to the immense range of dynamical behaviors that natural complex systems can exhibit. Where there is no symbolic resolution of an equation, numerical analysis involving algorithms (step-by-step recipes) can be used. It involves the discretization of space into cells, and time into steps (e.g. the heat equation under irregular boundary conditions).

Yet, most real-world systems do not even obey clear-cut macroscopic laws. For the majority of complex dynamical systems, global ODEs and spatial PDEs are no more possible because (a) there is no law of macroscopic quantities or global metrics that can represent the behavior of the system (no “ideal gas law” $PV = nRT$ for the embryo), (b) its myriad components are irregularly placed and mobile, requiring a non-Cartesian decomposition of space (the embryo does not conform to an x, y, z lattice), (c) there is a extreme degree of heterogeneity, requiring a segmentation into different classes of agents (the embryo is composed of many different cell types and behaviors), and (d) last but not least, most systems are highly adaptive, the topology and strength of the interactions depending on the short-term activity of the agents (e.g. cell migration) and the long-term fitness of the system in its environment (evolution). In short, there are no “Navier-Stokes” or “Maxwell” equations of embryogenesis.

A “multilevel” approach offers an alternative to the classical formalization. Instead of using a single equation ruling the behavior of a system at the macroscopic level, one can simulate the interactions of the elements at the microscopic level and observe the collective effect resulting from their aggregation. Traditionally in physics the process of deriving macroscopic laws from microscopic elementary interactions, which can be referred to as “coarse-graining” is possible with the averaging tools of *statistical mechanics* (e.g. a continuum description of the clock and wavefront model of somitogenesis derived from a chain of coupled oscillators [179]). These approximation or “mean-field” methods use probability theory to realize the transition from the discrete paradigm of local elementary behaviors to the continuous paradigm of macroscopic descriptions. However, they are constrained by strong assumptions such as the homogeneity of the elements. For example, the collective motion of cells in a developing embryo cannot be easily assimilated to fluid mechanics.

Once again, a widely popular solution comes from computer science under the terminology *agent-based modeling* (ABM). Historically, ABM originated from the need to model socio-economical systems too complex for analytical descriptions. Helped by the rise of computing technologies, it soon became a practical tool in many other scientific disciplines, such as ecology, biology and physics. Most of ABM is based on a combination of three types of topologies: fixed grids such as

square pixels, irregular networks with long-range connections, and 2D or 3D Euclidean space supporting mobile agents and metric-based interactions. Note that the continuous/agent-based dichotomy is more conceptual than practical, as many works mix both approaches at various degrees of compromise. As A.R.A. Anderson et al. put it [180]:

“Many mathematical models of biological process that consider space explicitly, fall into one of two categories: (i) continuum population models or (ii) discrete individual based models. Discrete, stochastic interactions between individual organisms cannot be captured by the continuum approach and likewise global population interactions cannot be captured by the discrete approach. In recent years a third category of models has emerged: hybrid models which allow modellers to exploit the advantages of both continuum and discrete models.”

2.2.3 Identifying Custom Laws at Each Level

In developmental biological systems, three levels of organization are considered: the sub-cellular level, in which the individual elements are molecules, the cellular level and the organism level. A “hard-liner” reductionist approach would skip the intermediate cellular level and describe the organism’s behavior in terms of molecular interactions. As mentioned above, this adventure would run into two major obstacles: not only the knowledge and theoretical models of molecular dynamics that we have today are not easily amenable to a numerical implementation, but also the available computational power, despite great advances, is still far from sufficient for a task of this magnitude. This is why the “emergent” approaches discussed above (statistical mechanics and/or agent-based modeling) have to be used, and *custom laws* have to be designed at each level. Even though cell-cell mechanical and chemical interactions are ultimately grounded in the physics of molecular interactions (covalent, ionic, hydrogen and electrostatic bonds), they seem to obey their own laws and “cell behavior ontology” on their phenomenological level. Naturally, the micro-to-macro transition is not unique as many different types of “higher-level” laws can “emerge” from the lower levels. The coarse-graining process is not only an approximation but also a selection process. Macroscopic behaviors described by higher-level laws are often focused on a specific set of features at that level of organization. Typically, the microscopic molecular level gives rise to macroscopic laws that can describe thermodynamic, metabolic, genetic, diffusive, or mechanical properties of the cell. Each law has its own domain of application.

Thus finding the laws that are particularly relevant to a model of biological development is not just a question of selecting an appropriate level of organization (cell, tissue, or sub-cellular components), but also what types of laws will be used and how they will eventually be coupled. The two major families of macroscopic properties that are considered in most developmental models are (a) biomechani-

cal properties and (b) molecular/genetic regulation and signaling properties. The MECAGEN project introduced in Section 1.2, to which our work contributes, constitutes a new type of morphogenesis modeling that integrates both (a) a representation of cellular and tissue biomechanics, and (b) networks of genetic and molecular interactions. While many models have been proposed to simulate the dynamics of genetic networks, none is directly involved in the regulation of physical properties of the cell. Yet, we believe that the key to understanding the morphogenesis and systemic properties of the organism lies in the *coupling* of cell biomechanics and networks of genetic and epigenetic regulation. It is about how (b) influences (a) through the production and modulation of the cytoskeleton, molecular motors, and cell adhesion, but also how (a) influences (b) through the transduction of mechanical stress. The modeling work should identify the appropriate level of schematization, i.e. capture essential causal relationships without going into fine molecular details. Each family is represented at the molecular level by different molecules:

- Molecules supporting the mechanics: The most important members of this family are the molecules that compose the cytoskeleton (actin, myosin, microtubules, microfilaments) and the molecules involved in adhesion, which attach the neighbor cells together (cadherins, integrins). In addition, although most other molecules are not directly responsible for the biomechanical properties of the cell, the whole molecular environment filling the cellular and extracellular volume creates important viscoelastic effects.
- Molecules supporting the molecular and genetic regulation and signaling: They are the well-known bases of the DNA and RNA, the amino acids composing the enzymes, and the whole transcription-translation-metabolic machinery in charge of creating, trafficking, and degrading metabolites and proteins.

These two categories of laws, mechanic and genetic, have been used and variously combined in the developmental models reviewed in Section 2.1, at one or several of the three main levels of the embryo. (a) Mechanical models have been proposed at the subcellular level, such as vertex-based junction models [181][182][147][145] and the Subcellular Element Model (ScEM) [183][184][148]; at the cell level, such as the Cellular Potts model [154] and off-lattice vertex models [185]; and at the organismal level, such as structural fluid mechanics inspired approaches [186][187]. (b) On the genetic and signaling side, source-sink diffusion models [123] and reaction-diffusion systems [100][113][122] at the tissue level; and gene regulation networks at the cell level, [188].

2.2.4 Summary Table

To conclude this review of models and principles of biological development, we propose below a summary table of some of the most exemplary works in chronological

order. For each of them, we specify schematically whether they follow a macroscopic approach (geometric, analytical, statistical) or agent-based methodology, and whether they include biomechanics, molecular/genetic dynamics, or both 2.8. We also indicate the type of cell behavior included, whether cells divide, the animal studied, whether the model is matched against real data and the simulation is in 3D 2.9.

2.3 Overview of the MECAGEN Modeling Principles

In the next three chapters, we explain the model that we have designed to contribute to the objective of the MECAGEN project.

- Chapter 3 describes how we express and calculate the mechanical interactions and behavioral properties of the cells. It presents a discrete-element model using one particle per cell driven by an overdamped equation of motion that can be summarily written $\lambda \vec{v}_i = \vec{F}_i^P + \vec{F}_i^A$, where \vec{v}_i is the velocity of one cell i , \vec{F}_i^P represents “passive” interaction forces controlling cell stiffness and adhesion, and \vec{F}_i^A represents “active” interaction forces, based on polarization axes, which create protrusive activity or apical constriction. These forces are calculated by summing over a neighborhood \mathcal{N}_i containing nearby cells that are in contact with cell i . This neighborhood is itself defined by a metric and topological criteria. \vec{F}_i^P is a *relaxation* force derived from an attractive/repulsive, elastic-like interaction potential. \vec{F}_i^A is a *behavioral* force moving the swarm configuration away from equilibrium.
- Chapter 4 deals with chemical signaling and gene regulatory networks. The goal is to briefly explain the principles of gene regulatory networks (GRNs), describe the components of GRN models, and give examples. In the present work, our particular objective was to design a simple and easily computable model of the molecular and genetic interactions that occur during development. Our model is articulated around three types of rules: rules driving the dynamics of intracellular gene/protein reactions, rules driving the dynamics of cellular secretion and transduction and rules driving the dynamics of extra-cellular reactions, transport and diffusion. They are expressed in a chemical kinetic framework by ODEs of the type $dp/dt = f(p, g, q, r)$, where p represents protein concentrations, g gene expression level, q external ligands, and r membrane receptors. Extracellular reactions, transport and diffusion of ligands are also taken into account via PDEs involving $\partial q/\partial t$ and fluxes $\vec{J} = -D\vec{\nabla}q$.
- Finally, the first steps toward building a complete morphogenetic platform integrating mechanics and genetics will be laid out in Chapter 5. A simplified “cell behavior ontology” (CBO) is proposed. It relates on the one hand *cell*

Year	Authors	Publication	Multi-agent / Macroscopic	Mechanics	Molecular and Genetic Regulation	Coupling
1917	D'Arcy Thompson	On Growth and Form	Macroscopic: describe geometrical transformation to relate the forms of animals	No	No	
1952	Alan Turing	The Chemical Basis of Morphogenesis	Macroscopic: (local)reaction-diffusion of morphogens, continuous	no	yes	
1969	Lewis Wolpert	Positional information and the spatial pattern of cellular differentiation.	Macroscopic: a bit qualitative, gradients induced field+polarity	no	yes	
1972	Hans Meinhardt & A. Gierer	A theory of biological formation	Macroscopic: pattern formation, continuous, reactions which combine self-activation and long-range inhibition	no	yes	
1990	Weliky, M. & Oster, G.	The mechanical basis of cell rearrangement. I. Epithelial morphogenesis during Fundulus epiboly.	particle based	yes, 2d polygons, internal pressure, damped motion equation	no	
1991	M. Weliky, S. Minsuk, R. Keller & G. Oster	Notochord morphogenesis in Xenopus laevis: simulation of cell behavior underlying tissue convergence and extension.	particle based	yes, 2d polygons, internal pressure, damped motion equation	no	
1991	Mjolsness, E., Sharp, D.H. & Reinitz, J.	A connectionist model of development.	Multi-agent	yes: but simple, user defined external forces	yes: reaction + diffusion towards neighbors	
1993	Glazier, J. & Graner, F..	Simulation of the differential adhesion driven rearrangement of biological cells.	first potts model	yes	no	
2001	Shapiro, B. & Mjolsness, E..	Developmental simulations with Cellerator	Multi-agent	Particle-based, elastic forces	yes, chemical reactions (ODEs)	yes
2001	Marée, A. & Hogeweg, P.,	How amoeboids self-organize into a fruiting body: multicellular coordination in Dictyostelium discoideum.	potts	yes	FitzHugh-Nagumo equations with piecewise linear "Pushchino kinetics" for the oscillating waves of cAMP	yes, add chemotactic potential in potts total energy from local cAMP gradient
2004	Dallon, J.C. & Othmer, H.G.,	How cellular movement determines the collective force generated by the Dictyostelium discoideum slug.	particle based	yes	no	
2005	Schaller, G. & Meyer-Hermann,	Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model.	Multi-agent, particle based one vertex/cell, topological	yes	yes and no: metabolic environment of the tumor	yes, define cell state
2007	K. Smith, Anna L. Langhans, Sara E. McLinden, Matthew A. Oberhardt, Karoly R. Kerec, Della P. D'Amato	Multiscale computational analysis of Xenopus laevis morphogenesis reveals key insights of systems-level behavior.	Multi-agent (Agent Based Model) netlogo...as terrible	On-lattice but not potts, agent can overlap on 2D pixel to emulate 3D. Cell has 3^3 surface.	yes, intracellular reaction	strange...
2007	D.W. DeSimone, J.A. Papin and S.M. Peirce	Cell adhesion and cortex contractility determine cell patterning in the Drosophila retina	potts: add cell cortical contraction to adhesion-lead surface increase	yes	no	
2008	M. Krieg, Y. Arboleda-Estudillo, P.-H. Puech, J. Käfer, F. Graner, D. J. Müller and C.-P. Heisenberg	Tensile forces govern germ-layer organization in zebrafish	potts	yes	no	
2008	Honda, H. et al.,	Computer simulation of emerging asymmetry in the mouse blastocyst.	particle based: a vertex / neighboring position. Damped equation of motion. Surface energy, volume conservation	yes	no	
2008	Rejniak, K.A. & Anderson, A.R.A.	A computational study of the development of epithelial acini: I. Sufficient conditions for the formation of a hollow structure.	mechanics both particle-based and continuous (navier-stokes) for the fluid	yes. Elastic particle for cell membrane + viscous incompressible fluid inside and outside the cell	no, cell type is specified by adhoc rule	
2010	Scott Christley, Briana Lee, Xing Dai, Qing Nie	Integrative multicellular biological modeling: a case study of 3D epithelial development using GPU algorithms	ScEM	yes, scem + custom adhesion forces	ODE set for Notch Delta intra and inter cell signaling	Cell cycle is coupled with the molecular and genetic regulation
2011	S. Sandersius, M. Chuai, C. J. Weijer, T. J. Newman	Correlating Cell Behavior with Tissue Topology in Embryonic Epithelia	ScEM, particle based: multiple vertex / cell. No difference between cell membrane and innercell particle in structure. Morse potential. Not topological	yes	no	substance concentration levels induce cell growth. Mitosis if threshold is reached (size threshold)
2011	Susan D. Hester*, Julio M. Belmonte, J. Scott Gens, Sherry G. Clendenon, James A. Glazier	A Multi-cell, Multi-scale Model of Vertebrate Segmentation and Somite Formation.	Potts + genetic regulation (using CompCell3D)	yes	yes (intra + cell-cell communication + diffusion)	boolean cell-type determination network
2011	S A Sandersius, M Chuai, C J Weijer and T J Newman	A "chemotactic dipole" mechanism for large-scale vortex motion during primitive streak formation in the chick embryo.	ScEM, particle based: multiple vertex / cell. No difference between cell membrane and innercell particle in structure. Morse	yes, include mechanotaxis for inducing migration axis	no	a linear diffusion with poison-noise degradation field induce polarization axis for chemotactic behavior

Figure 2.8: Summary table of the developmental models and principles reviewed in Section 2. (Part I)

Year	Authors	Publication	Cell behavior	Cell division	Animals studied	data matching	3D
1917	D'Arcy Thompson	On Growth and Form	none	no	various	no	no
1952	Alan Turing	The Chemical Basis of Morphogenesis		no	tentacle of hydra	no	no
1969	Lewis Wolpert	Positional information and the spatial pattern of cellular differentiation.	growth of tissue	yes but qualitative...	early dev. Of sea urchin, regeneration of hydroids	no	no
1972	Hans Meinhardt & A. Gierer	A theory of biological formation	polarization in a gradient field	no	hydra	no	no
1990	Wolpert, M. & Oster, G.	The mechanical basis of cell rearrangement. I. Epithelial morphogenesis during Fundulus epiboly.	migration towards vegetal pole (epiboly)	no	fundulus epiboly	no	no, 2D on a sphere
1991	M. Weliky, S. Minsuk, R. Keller & G. Oster	Notochord morphogenesis in <i>Xenopus laevis</i> : simulation of cell behavior underlying tissue convergence and extension.	intercalation/polarization	no	xenopus laevis	no	no
1991	Mjolsness, E., Sharp, D.H. & Reinitz, J.	A connectionist model of development.	differentiation	Yes , grammar rules	blastoderm of drosophila	no	no
1993	Glazier, J. & Graner, F.	Simulation of the differential adhesion driven rearrangement of biological cells.	cell sorting, cell dispersal			no	no
2001	Shapiro, B. & Mjolsness, E.	Developmental simulations with Cellerator	growth, death	yes			yes
2001	Marée, A. & Hogeweg, P.,	How amoeboids self-organize into a fruiting body: multicellular coordination in Dictyostelium discoideum.	chemotaxis	no	dictyostelium discoideum	no	no
2004	Dallon, J.C. & Othmer, H.G.,	How cellular movement determines the collective force generated by the Dictyostelium discoideum slug.	collective migration	no	dictyostelium discoideum	no	yes
2005	Schaller, G. & Meyer-Hermann,	Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model.	mitosis, apoptosis (tissue growth)	Yes , grammar rules	tumor cells	yes	no
2007	Scott H Robertson, Chris K Smith, Anna L Langhans, Sara E McLinden, Matthew A Oberhardt, Karoly R Jakab, Bette Dzamba	Multiscale computational analysis of <i>Xenopus laevis</i> morphogenesis reveals key insights of systems-level behavior.	migration on fibronectin matrix	no, 53 cells	mesendoderm migration in <i>Xenopus laevis</i>	no	2D
2007	D.W. DeSimone, J.A. Papin and S.M. Peirce	Cell adhesion and cortex contractility determine cell patterning in the Drosophila retina	Converged cell shapes according to initial mechanical studies	no	drosophila retina	automated but nice visual compariso	no
2008	M. Krieg, Y. Arboleda-Estudillo, P.-H. Puech, J. Käfer, F. Graner, D. J. Müller and C.-P. Heisenberg	Tensile forces govern germ-layer organization in zebrafish	cell sorting according to various mechanical parameters	no	zebrafish	no	no
2008	Honda, H. et al.,	Computer simulation of emerging asymmetry in the mouse blastocyst.	mechanical evolution of cell shapes according to initial mechanical studies	no, constant 40 cells	mouse blastocyst	no	yes
2008	Rejniak, K.A. & Anderson, A.R.A.	A computational study of the development of epithelial acini: I. Sufficient conditions for the formation of a hollow structure.	Polarization (epithelial), apoptosis, growth	yes	epithelial acini	no	no
2010	Scott Christley, Briana Lee, Xing Dai, Qing Nie	Integrative multicellular biological modeling: a case study of 3D epidermal development using GPU algorithms.	growth	yes	mammalian epidermis		yes
2011	S. Sandersius, M. Chuai, C. J. Weijer, T. J. Newman	Correlating Cell Behavior with Tissue Topology in Embryonic Epithelia	Cell growth	yes	hydra, drosophila, xenopus, chick number of neighbor in epithelial cells	yes, no fitness	no
2011	Susan D. Hester*, Julio M. Belmonte, J. Scott Gens, Sherry G. Clendenon, James A. Glazier	A Multi-cell, Multi-scale Model of Vertebrate Segmentation and Somite Formation.	cell motility controloed by a parameter	no, add cells at the posterior end of the presomitic mesoderm	vertebrate somitogenesis	no	no
2011	S A Sandersius, M Chuai, C J Weijer and T J Newman	A "chemotactic dipole" mechanism for large-scale vortex motion during primitive streak formation in the chick embryo	chemotaxis, mechanotaxis	No, 1200 cells constant	primitive streak formation in the chick embryo	qualitative	no

Figure 2.9: Summary table of the developmental models and principles reviewed in Section 2. (Part II)

states, determined by the concentrations of certain proteins, and on the other hand *cell behaviors*, determined by biomechanical parameters.

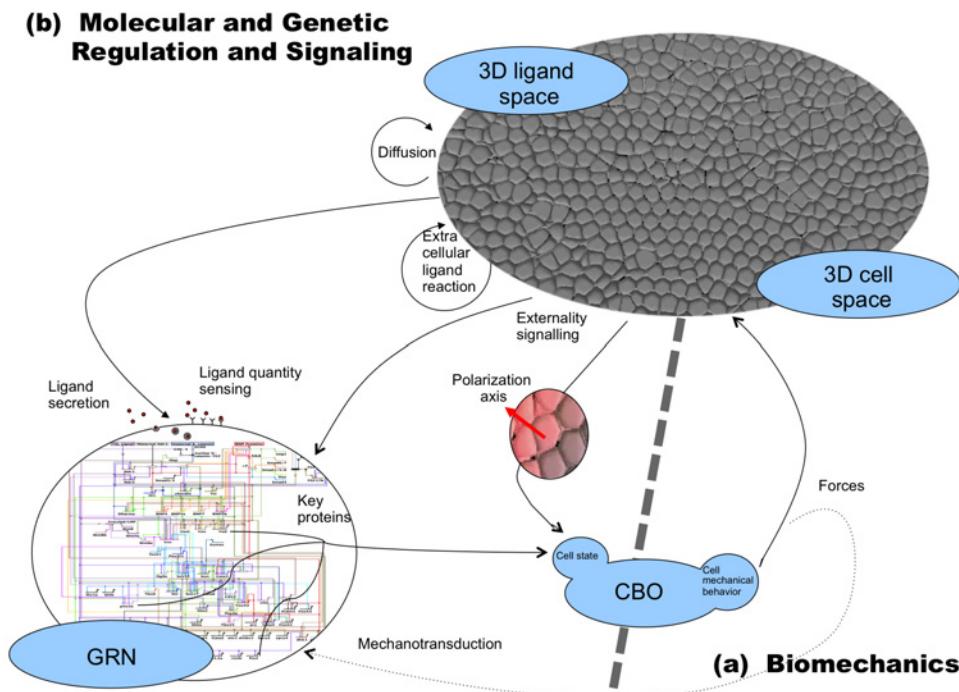


Figure 2.10: General diagram of the MECAGEN project.
 a: Biomechanical processes, including a “passive” and “active” forces (see summary of Chapter 3 above). b: Molecular/genetic regulation and signaling processes, including gene regulation networks, intracellular protein concentrations, extracellular ligands secretion, diffusion and binding.

CHAPTER 3

MECAGEN Model of Cell Biomechanics

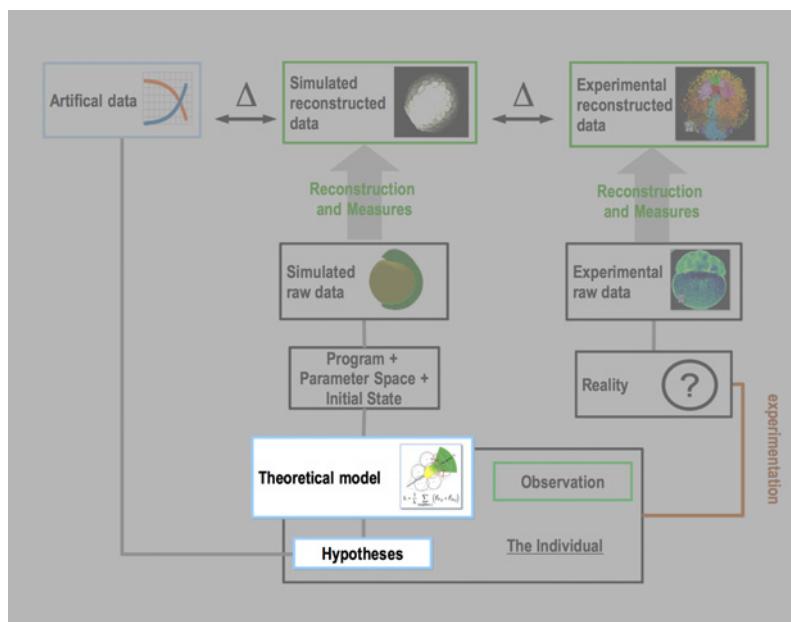


Figure 3.1: Situation of Chapter 3 in the methodological workflow.

In this first chapter of three that are dedicated to the model, we explain the choices that we made about representing and calculating the mechanical interactions and behavioral properties of the cells (Fig. 3.1). Chemical signaling and internal gene expression will be the topic of Chapter 4 (albeit in less detail), while the overall plan of building a complete morphogenetic platform integrating mechanics and genetics will be exposed in Chapter 5. Altogether, our agent-based modeling and simulation endeavors set the stage for an “*in silico*” embryo, i.e. a virtual test object that can be manipulated and experimented upon (like putty or clay) in ways impossible with a real embryo. The goal is of course to better understand embryogenesis, and how the particular zebrafish study represents only one specific parametric instance among a much broader class of generic phenomena. Naturally, this virtual embryo, although deprived of “life”, still behaves very much like a living entity, ani-

mated by the processing units and execution threads of the computers that calculate it. It is a form of “active matter” composed of a great number of self-propelled particles (cells), which carry instructions (rules and local data) diverse enough to equip them with a rich repertoire of potential “decisions” and actions that are function of what they sense from their surroundings. Note that the global environment here will be reduced to a simplified physical world, where gravity doesn’t play any role, so that the only forces maintaining the cohesion of the embryo will be found among the cells.

3.1 State of the Art in Biomechanical Modeling

Before describing the cell mechanical model that we have designed (in Section 3.2), we present here a brief review of the scientific literature. We start with a sampler of different families of models in Section 3.1.1, from continuous to discrete spaces, of geometrical or physical inspiration, then we focus on the specifics of *particle-based methods* in Section 3.1.2, which will serve as a basis for our own approach.

3.1.1 The Many Viewpoints on “Soft Matter”: From Continuous-Space to Discrete-Element Models

By explicitly introducing the concept of *Entwicklungsmechanik* in the 1880’s, Wilhelm Roux established the importance of mechanics in the study of developmental systems. Since then, a great number and diversity of theoretical models of biomechanics have been proposed at several levels of abstraction—and speculation. The unfolding and expansion of the embryo, an endless source of fascination, has been examined from vastly different viewpoints from all corners of science and engineering. Depending on the researchers’ background and their focus of interest, embryogenesis (formerly “embryology”) has been assimilated to differential geometry, pattern formation, fluid dynamics, material physics, systems architecture, cellular automata or collective motion, among many other fields and disciplines. The growth and shaping of cells, tissues and organs have been variously compared to manifolds, balloons, tensegrity structures, bubbles, swarms, and so on. Truth is, the extraordinarily rich biological material that composes a multicellular organism is all of the above—and much more. At the emergent, phenomenological scale of observation, it seems to obey nothing like the common-sense physical laws and invariant geometrical properties of mundane, human-scale artefacts. Nonetheless, these attempts are all useful to help us capture a piece of this puzzle and bring us closer to a more accurate and complete picture.

We propose here a sampling of different broad families of theoretical models that have tried to describe the *spatial dynamics* of the growing embryo by various means. Their common purpose is to propose a mechanical explanation that should be necessary and sufficient to generate the observed kinematic behavior of the cell

tissue. Again, the diversity of these approaches is due to the extremely elusive nature of multicellular systems compared to the traditional objects studied by physics. Classical physics has developed a framework suited to fixed objects presenting high spatio-temporal regularities. In contrast, the living matter is riddled with heterogeneity, irregularities and ceaseless internal dynamicity and adaptivity. Local cell behavior and global tissue properties can change rapidly as they rest upon a molecular structure in constant flux and state of self-reorganization. This heterogeneity of behavior also gave rise to a heterogeneity of models, some more specialized than others in certain aspects of cell dynamics, and unfortunately not always immediately compatible with each other (for example, the concept of “interfacial tension” is not a clear contributor to the phenomenon of “protrusion”; see below).

In general, cells behave neither like a solid nor like a fluid—rather, they behave like both at the same time. This is why cell tissue is sometimes described as *soft matter*. Another fundamental notion to keep in mind when describing the behavior of cells is their ability to perpetually convert stored energy into mechanical work, which puts them into permanent “out-of-equilibrium” states, far from the descriptions of classical mechanics and thermodynamics. This notion is at the foundation of what is called today *active matter* in modern physics, and is applied to various levels of organization, from macromolecules to cells to flocking birds [189]. Qualifiers commonly used for these objects are “self-x” (e.g. self-propelled, self-healing, etc.) or “spontaneous”, since they convey the idea of an “agent” capable of autonomous actions independent from external influences. In egg-contained embryos, the source of energy necessary for the motion and behavior of the cells is found in the yolk. In mammals, the energy is provided by the mother essentially in the form of nutrients, not direct mechanical action.

In this section, we distinguish among the different approaches to cell dynamics based on whether they adopt a macroscopic viewpoint relying on continuous space or a microscopic viewpoint relying on discrete elements (Fig. 3.2):

- Macroscopic viewpoint: *continuous-space* descriptions of the embryo set behavioral laws directly at the global tissular level without explicit underlying cellular or molecular components. They are generally based on macroscopic (spatial) differential equations, which have the advantage of compactness, as they offer an inclusive representation of development in one or a few formulas. On the other hand, the main disadvantage of grand formalisms can also be too much generality and vagueness, with a consequent lack of specific operational tools. In any case, cell tissue in this paradigm is construed as a continuous mass, equivalent to an infinity of infinitesimal points. We distinguish here two main perspectives, both forms of mathematical biology: solid mechanics and fluid mechanics.
- Microscopic viewpoint: *discrete-element* descriptions of biological tissue consider the cells (and sometimes the subcellular structures and molecules) that

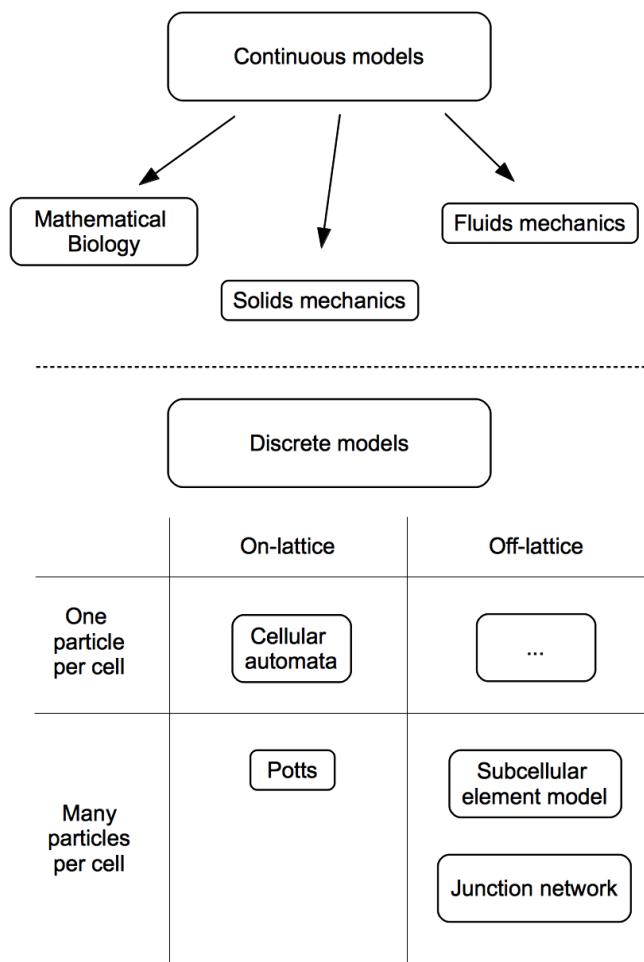


Figure 3.2: The schematic categorization of the models adopted in this review.

constitute it as autonomously acting components. It is their collective behavior that determines the mechanical properties of the tissue at the emergent level. As they do not always coincide one-to-one with cells (sometimes sub-cellular or supracellular), the elements of discrete models are alternatively called “particles”. Their properties generally include spatial coordinates and geometric properties, and optionally mechanical/physical properties, too. We distinguish here four groups, depending on whether particles are confined to a discrete grid or not, and cells are made of several particles or only one.

Continuous Models

Mathematical Biology

Among James Murray’s classical works in mathematical biology [119], [95], Murray and Oster propose a specific model of biological pattern formation [190], which describes the behavior of a tissue comprising mesenchymal cells. In this model, the temporal rate of cell density is related to the flux of cells and the mitotic rate. The flux function takes into account different transport effects such as convection and random dispersal. The following two sections illustrate this continuum mechanics paradigm, which treats the embryo as “bulk matter”, and is broadly divided into solid mechanics and fluid mechanics. A distinctive feature of solids and fluids is their unequal ability to resist to the action of a shearing force tangential to the surface. Both fields distinguish different types of bulk behavior and both have been applied to biological matter and multicellular tissue. In the case of fluids, stresses are linked to velocity fields through a continuity equation and conservation laws. In the case of solids, stresses are linked to a deformation tensor (strain) [186].

Solid Mechanics

Solid mechanics focuses on the study of the response of solid matter to the action of forces and/or thermal loading. The most important theoretical tool used in solid mechanics is the *tensor*: it is a 3D matrix that represents an ellipsoid whose axes convey the principal directions and amplitude of the gradient of some quantity in a local neighborhood. Typical relevant quantities are the strain and the stress. The strain expresses the deformation of the solid, and is deduced from the displacement of the solid. The stress is a measure of the internal forces acting inside the solid as a reaction to an external action, and its dimension is force per unit area (Fig. 3.3). The relation between the stress tensor and the strain tensor establishes the mechanical properties of the matter.

These mechanical properties can be of various types, such as linear/nonlinear elasticity (hyperelasticity), plasticity and elastoplasticity, or viscoelasticity:

- **Linear elasticity** refers to the ability of a material to recover its original shape

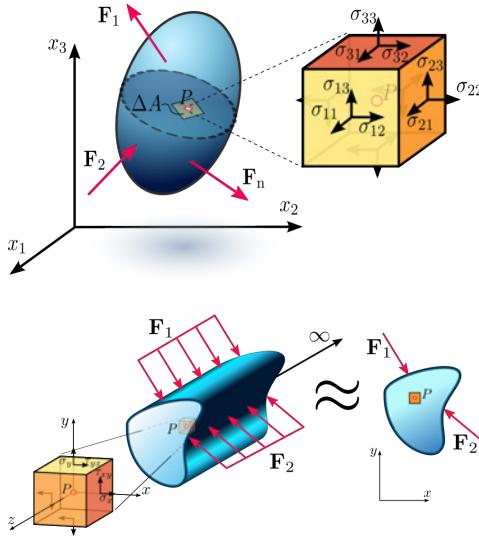


Figure 3.3: Stress and Strain. Left. Stress from a loaded deformable material body assumed as a continuum from Wikipedia. Right. Plane strain state in a continuum from Wikipedia.

after external forces, causing relatively small deformations, are removed. It is deemed “linear” when the material deformation is proportional to the applied load (e.g., the classical Hooke’s law $x = F/k$ of springs). An example of elastic model in biological development can be found in [191]: here, the meristem of *Arabidopsis* is represented by two layers of polyhedral cells interacting via their walls, which are considered isotropic elastic material. Outer walls are made stronger than inner walls due to an elevated elastic modulus.

- **Nonlinear elasticity:** nonlinear elasticity, however, is more appropriate to the description of most observed material behavior, especially biological matter. For example, in [192] a 3D finite element model of ventral furrow invagination in the *Drosophila melanogaster* embryo makes use of neo-Hookean “hyperelastic” material.
- **Plasticity** refers to materials that can undergo permanent deformation after forces have been applied to them (e.g. metals or soils). In combination with elasticity, “elastoplasticity” can also describe the peculiar behavior of certain biological tissues. For example, in [186], embryonic sheets in the chick are viewed as a very soft solid that can deform and yield. In a first approximation, the elastic deformation of the blastodisc sheets is calculated and further assumptions lead to the conclusion that the deformation is in fact proportional to the gradient of pressure.
- **Viscoelasticity** qualifies solid materials that exhibit fluid-like properties,

such as certain types of plastics. For example, [190] proposes a model of epithelium pattern through a viscoelastic field equation system, in which the contraction is triggered by a calcium pulse. Another quantitative study of living embryonic tissues [193] presents an experimental technique that allows the compression and relaxation of cellular aggregates. Here, a generalized Kelvin-body model of viscoelasticity is used.

Fluid Mechanics

Fluid mechanics focuses on the study of liquids, gases and plasmas, and their response to forces applied to them. As a part of continuum mechanics, fluid mechanics assumes that the properties of fluid density, pressure, temperature and velocity are continuous. The fundamental hypotheses of fluid mechanics are the conservation of mass (rate of mass change in an elementary volume equals rate of incoming mass minus rate of outgoing mass), conservation of energy, conservation of momentum, and continuity. Concerning the last hypothesis, the “continuity equation” ensures that mass, energy and momentum are also *locally* conserved as they can only move in a continuous flow. The principal equation ruling fluid behavior is the Navier-Stokes equation. It describes the motion of the fluid as a function of the balance of forces acting on it. Its general form for the conservation of momentum reads

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = \frac{1}{\rho} \nabla \mathbb{P} + \frac{1}{\rho} \mathbf{F}$$

where \mathbf{v} is the velocity vector, ρ is the density of the fluid, \mathbb{P} is a tensor which describes the normal and tangential stresses applied on a fluid particle, \mathbf{F} is a body force vector (force per unit volume). Additional conditions on the fluid proprieties can lead to a solvable set of equations. If the fluid is incompressible, the divergence of the stress term $\nabla \mathbb{P}$ is replaced by the sum of a pressure gradient $-\nabla p$ for the normal component of the stress and, for a Newtonian fluid, a dynamic viscosity constant μ times the Laplacian of the velocity $\nabla^2 \mathbf{v}$ expresses the tangential component of the stress.

Unfortunately, in biological material, the Newtonian assumption is hardly satisfying as the viscosity model is not linear. The use of more complicated non-linear stress tensor is required and under particular external conditions, holes may appear in the continuum. However, a simplifying assumption adapted to biological material can be made as the inertial forces (LHS of the Navier-Stokes equation) are negligible compared to the forces of viscosity (RHS of Navier-Stokes). The Stokes equation expresses this assumption:

$$\nabla \mathbb{P} + \mathbf{F} = 0$$

Only the simplest forms of the Navier-Stokes equation are analytically solvable (e.g. Poiseuille flow, Couette flow, Taylor-Green vortex, etc.), more complex con-

ditions requiring the use of computer simulations also called Computational Fluid Dynamics (CFD).

Among examples of developmental models that make use of the Navier-Stokes equation, [194] proposes a model of branching morphogenesis involving both an epithelial tissue and a mesenchymal tissue to behave like Stokes fluids with different viscosities separated by an interface. In a hydrodynamic simulation of multicellular embryo invagination [195], the authors successfully couple the behavior of a viscous incompressible fluid ruled by Stokes equation with the movement of a membrane of cells during epithelial invagination. Another model of the chick embryo based on a discontinuity of tissue flow leading to the formation of the cephalic fold [196] also uses Stokes fluid to model folding processes.

Discrete Models

In this category, models can be contrasted depending whether they also discretize the space (which contains the discrete elements) or not. The spatial coordinates of the elements may either belong to a grid, i.e. integer multiples of given elementary distances, or the entire Euclidean space (\mathbb{R}^3), i.e. unconstrained real values. We call the first type of approach “on-lattice”, and the second type “off-lattice”. Another distinction, orthogonal to the previous one, can be made across the number of particles used to represent a cell, whereas it is one or many. The combination of these two dimensions gives rise in principle to four categories (Fig. 3.2).

On-Lattice Models with One Particle per Cell

This category is represented by *cellular automata* (CA), which assimilate a biological cell to one grid cell (“pixel” in 2D, “voxel” in 3D). A cell’s neighborhood can be defined in various ways, whether comprising only the immediate “orthogonal” neighbors (4 in 2D, 6 in 3D) or including the “diagonal” neighbors, too (for a total of 8 in 2D, 26 in 3D), sometimes even taking into account farther away neighbors. Generally, the fate of each grid cell at each time step (i.e. its change of internal state) is decided by a rule that takes into account the combined internal states of all its neighbors. A great number of biological models in various domains have been using a CA framework [197]. Typical cases of CA-based models include tumor growth and invasion [198], and cell migration [199]. There, cells can for example migrate and interact on a 2D or 3D grid (south, west, north, east or no move) with probabilities based on continuous values derived from a random walk. Generally, the modeler must resolve spatial conflicts by defining some occupancy policy, i.e. whether two cells may be located at the same grid point at the same time or not. In [198], cells are allowed to proliferate and move only toward empty neighboring locations. Other typical cases feature models of plant growth [200].

On-Lattice Models with Multiple Particles per Cell

This category is essentially represented by the *Cellular Potts Model*, a successful paradigm for modeling multicellular assembly. It was introduced by F. Graner and J. Glazier in 1992 and 1993 [153], [154] as an adaptation of the Potts model to reproduce the cell sorting behavior described by Steinberg [154]. Historically, the Potts model was derived from the Ising model (itself invented by Wilhelm Lenz in 1920), which describes the phase transition of the magnetic dipole moments of atomic spins arranged on a lattice and taking binary values (+1 or -1). In particular, the Ising model gives a characterization of the transition between the ferromagnetic and paramagnetic phases. Spins only interact with their four immediate neighbors on the lattice (local principle). Coupling between spins is quantified by an energy function called the Hamiltonian, equal to the sum over every atom of the coupling factors between an atom and its neighbors: $H = -\sum_{ij} J_{ij} s_i s_j$, where aligned spins make a lower energy contribution than anti-aligned spins. In 1943, Ashkin and Teller added an update dynamics to the model based on a Monte-Carlo algorithm and the Boltzmann probability (later called a Metropolis algorithm).

In 1952, Renfrey Potts extended the Ising model by allowing the spin to take multiples values. With a Metropolis algorithm, the Potts became a kinetic simulation and allowed the study of various systems from metal grains to foam [201]. The *cellular* Potts model applies the model to cells by stating that a cell is represented by an ensemble of neighboring grid points and the “spin” of each grid point is the id of the cell (usually represented by different colors; Fig. 3.4). Its first version used an Hamiltonian with a coupling term representing the differential adhesion coefficients according to cell types and a volume conservation constraint penalizing cell size variation. The cellular Potts model was later extended in collaboration with Paulien Hogeweg and colleagues [202][203][204], in particular producing a successful model of *Dictyostelium* aggregation and growth [205], to allow the integration of various cell behaviors: cell growth, cell division, cell migration. A modulation term was also introduced in the Hamiltonian to take into account the rigidity of the cell through a surface conservation law. To acknowledge these improvements, the model was renamed Glazier-Graner-Hogeweg (GGH). An open-source modeling environment called CompuCell3D [168][169] was also developed based on this model, and has been gathering a growing community of users. Since then, many works following the cellular Potts approach have been published in domains as diverse as development and cancer growth [170], cell differentiation and migration [206], cell sorting by differential adhesion [159], vertebrate segmentation and somite formation [207][208], and vasculogenesis based on chemoattractant gradients [209].

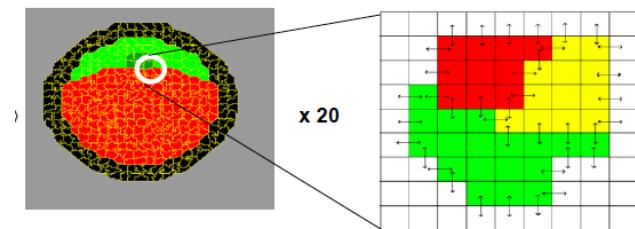


Figure 3.4: Cellular Potts Model on-lattice space.

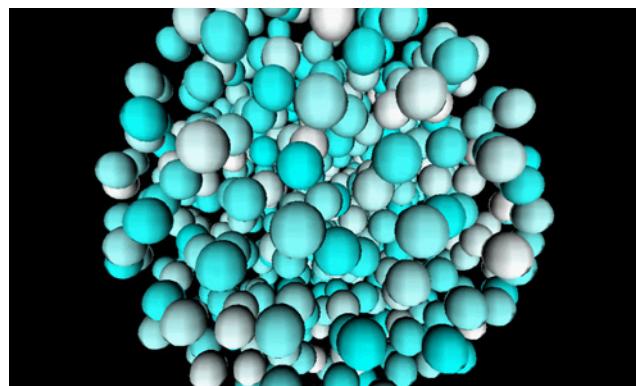


Figure 3.5: One of our early experiments of single-particle cell models exploring the biomechanics of the MECAGEN project. Each cell is represented by a small sphere and interacts with its nearest neighbors in a Delaunay tessellation through elastic forces (see Section 3.2).

Off-Lattice Models with One Particle per Cell

In physics, the field of “soft matter” concerns the study of macromolecules, such as polymers, colloids, or gels, which generally have non-trivial emergent properties that cannot easily be predicted based purely on their atomic composition. Since the focus of soft matter physics is subcellular phenomena, some models make the choice a higher level of abstraction at which the mechanical properties of a cell and assemblies of cells can be more easily accounted for. Here, the system is typically composed of a 2D or 3D irregular network, whose vertices represent cells and edges interactions (e.g. adhesion) between cells. This is the framework that we eventually adopted in our own study (Fig. 3.5 and Movie 3.1; see Section 3.2), after having experimented with multiparticle systems (see next). An early example of single-particle cell model in a 2D culture was proposed by Dirk Drasdo in 1995 [210]: it included visco-elastic and classical attraction-repulsion potentials, and was simulated by a Monte-Carlo method. Another 3D model of cell movement in multicellular systems, with net propulsion force (without reaction), surface approximation, and ellipsoidal particles can be found in [211]. Meyer-Hermann et al. have designed an agent-based and force-based “Delaunay-Object-Dynamics” framework in which neighborhoods are derived from a weighted Delaunay triangulation [185] optionally calculated in parallel [212], and volumes are conserved by a pressure force. This is very similar to what we use in our own model as it will be explained in Section 3.2. They apply it to models of avascular tumor growth and morphogenesis of lymphoid follicles [213], epidermal homeostasis [214], and cell migration with contractile ring along the migration axis [215]. Multicellular migration can also be simulated by combining elastic forces and migration forces [216], assuming that a qualitative distinction can be schematically made between “passive” and “active” forces as it is the case in our study (Section 3.2).

Off-Lattice Models with Multiple Particles per Cell

More detailed and complex models have chosen to preserve the fine grain of subcellular particles, as in the cellular Potts paradigm but in quasi-continuous space without a discrete grid support. Two major categories can be identified in this family: vertex-based junction networks and the Subcellular Element Model (ScEM). It is also worth noting *tensegrity* structures, composed of compression-bearing rigid elements (rods) isolated from each other and connected by a network of tensioned cables (springs)—an engineering and architectural concept used in abstract models of the cytoskeleton, notably by Donald Ingber [217][218]. Although the properties of one tensegrity structure can well represent the flexibility and resilience of a single cell, it is difficult to aggregate and link several of these together, thus there are virtually no models of *multicellular* tensegrity-based tissues or organisms.

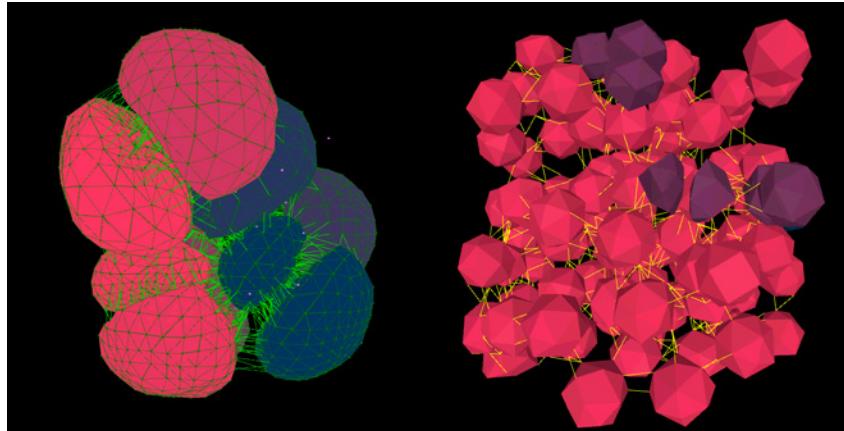


Figure 3.6: Two of our early experiments of multiparticle cell models exploring the biomechanics of the MECAGEN project. Left: Each cell is a sac filled with a great number of subcellular elements, appearing as vertices of a triangular mesh on the cell membrane. Green edges schematize intercellular adhesion forces (Movie 3.2). Right: Similar experiment with simpler empty polyhedral cells.

Vertex-based junction networks

In this category, cells are represented only by their membrane, which is generally formatted as a triangular mesh of vertices, but are not filled with other elements. The advantage is an explicit *surface* description of cell shapes, however it makes large-scale multicellular motion and rearrangement more difficult to simulate. A precursor computer simulation based on this idea was due to Odell et al. in 1981 [219]. It was designed as a 2D model of epithelial folding and invagination to test hypotheses about the mechanical properties of cells. Cytoskeletal microfilaments form a cuboidal structure whose apical facet is contractile. Neighbor cells are laterally connected so that the contraction of the apical facet creates an invagination in the tissue. The contraction is triggered by a mechanical signal (stretching) that propagates along the aligned cells. Another notable attempt to model multicellular development as a growing 2D tiling of polygonal cells (possibly projected on a sphere) was presented by Weliky et al. in 1990 [182]. It is one of the rare works to include cell protrusion and intercalation mechanisms (by “extension” of one node and “inhibition” of the others) with a rule to merge vertices when cell shapes change. They illustrate their system on the rearrangement of the enveloping layer during epiboly of the *Fundulus*[181], and convergence-extension in the *Xenopus* notochord [182]. The same type of network can support simpler models of cell division patterns in a growing epithelial packing, where cell rearrangement is only due to growth and not to protrusion or migration [147][145]. Vertex-based junction networks have also

been variously employed in meristematic growth in plants (*Arabidopsis*[191]) based on a global energy term including internal pressure and wall stiffness, the mouse blastocyst (a spherical embryo with an inner cavity) [220], the ascidian endoderm invagination by a two-step mechanism (apical contraction, then collared rounding of the cells) and multipoint curved edges between vertices [221], and the wing imaginal disc of *Drosophila*[222].

Subcellular Element Model (ScEM)

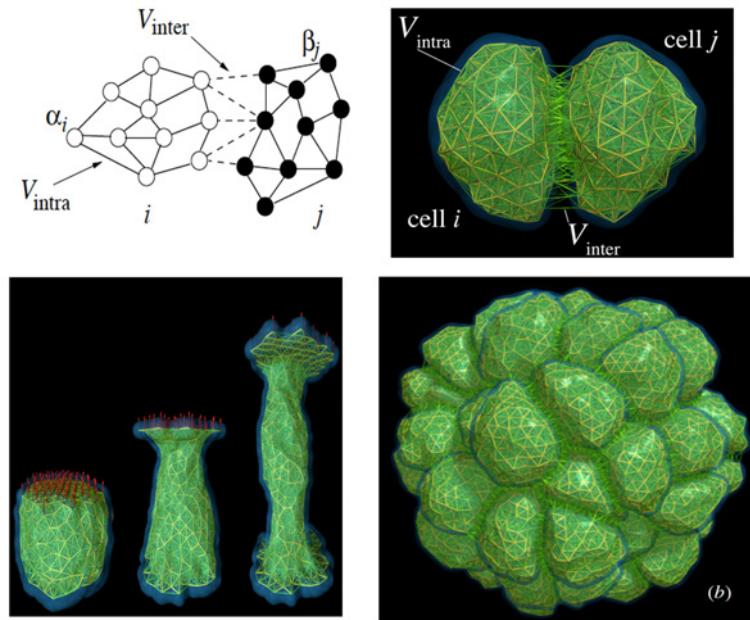


Figure 3.7: Illustrations of the Subcellular Element Model (ScEM). Top left: Schematic diagram showing two cells, i and j , and a subset of the intra- (solid lines) and intercellular (dashed lines) interactions between their sub-elements (open circles in i , filled circles in j). Image and caption adapted from [183]. Top Right: Visualization of two interacting cells simulated using the ScEM and comprising about 500 elements each. Image and caption from [223]. Bottom left: Visualization of an ScEM simulation of a single cell being stretched by a constant unilateral applied force. Image and caption from [223]. Bottom right: A spherical cluster of about 110 cells with about 370 elements per cell. Image and caption from [224].

ScEM was introduced by Newman in 2005 [183]. In this model, 3D cells are composed of multiple subcellular elements (typically a few hundred for a small number of cells, and a few dozen for thousands of cells [184]). The interaction rules among these elements are similar to the cell-cell interactions of the single-particle models

at a lower scale. The interest of this approach is mainly that cell shapes and their dynamics are emergent properties of the underlying particles, both on the surface and in the interior of the cells (see review in [223]). The visco-elastic properties of ScEM were tested in rheological experiments by applying stress on putty-like bulk material [224] (Fig. 3.7, bottom left), then used in a study of collective migration in the primitive streak formation in the chick embryo [184], and epithelial behavior in various species [148].

Review literature

- John L. Semple, Nicholis Woolridge, and Charles J. Lumsden, 2005. In vitro, in vivo, in silico: computational systems in tissue engineering and regenerative medicine. *Tissue engineering*, 11(3-4), pp.341–356. [225]
- David Gonzalez-Rodriguez, Karine Guevorkian, Stéphane Douezan, Francoise Brochard-Wyart, 2012. Soft Matter Models of Developing Tissues and Tumors. *Science*, 338(6109), pp.910–917. [226]
- Matthew A. Wyczalkowski, Zi Chen, Benjamen A. Filas, Victor D. Varner and Larry A. Taber, 2012. Computational models for mechanics of morphogenesis. *Birth Defects Research Part C: Embryo Today: Reviews*, 96(2), pp.132–152. [227]
- Andreea Robua, Roxana Aldeab, Oana Munteanuc, Monica Neaguc, Lacramioara Stoicu-Tivadar, Adrian Neagu, 2012. Computer simulations of in vitro morphogenesis. *Biosystems*, 109(3), pp.430–443. [228]

3.1.2 Particle-Based Physics

In particle-based physics models, each element of matter is represented by a particle. A great number of particles put together represent a higher level entity, such as a cell. The mechanical properties of the system are given by the parameters of the interaction potential between neighboring particles. When designing a particle-based model, two main question must be addressed: (a) How are the space and the objects it contains divided into particles (spatial neighborhoods)? (b) How is the motion of particles related to their mechanical properties (attraction-repulsion potential)?

Spatial Neighborhoods

Equations ruling a particle system's dynamics are based on a double sum over all particles i and their neighbors $j \in \mathcal{N}_i$. This section describes the different types of neighborhoods that can be defined. In **off-lattice** simulations, particles are free to move to any coordinates in the 3D Euclidian space \mathbb{R}^3 . Thus neighborhoods are

not implicit as they are on a grid, and an algorithm is needed to explicitly identify neighboring particles. We denote by P the swarm of n particles i at positions $\vec{X}_i = (x_i, y_i, z_i)$.

All to All Neighborhoods

In the simplest, but most computationally expensive case every particle is connected to all the others, creating $n(n - 1)/2$ bidirectional links in the swarm P .

$$\mathcal{N}_{ij}^{\text{all to all}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

Metric (Distance-Based) Neighborhoods

In the metric case, the intensity of the interaction between particles decreases when their distance increases. At some point, it may even be neglected with respect to interactions with closer particles. Thus by choosing a threshold, a definition can be that two particles are considered neighbors if and only if they interact with an intensity above that threshold. In other terms, it is as if each particle generated a local field of interaction potential with limited scope. In this case, the neighborhood matrix can be much sparser, for example:

$$\mathcal{N}_{ij}^{\text{metric}} = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & \cdots & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & \cdots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \cdots & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & \cdots & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

In its simplest form, this locality principle may be captured by a intrinsic size parameter s , applied to all particles of P : the local domain is then defined by a

sphere of radius $f(s)$. In Euclidean space, it means that

$$\|\vec{X}_i - \vec{X}_j\| \leq f(s)$$

If particles have spheroidal shapes, an intrinsic radius parameter r_i can be associated to each particle i , in which case the metric criteria reads

$$\|\vec{X}_i - \vec{X}_j\| \leq f(r_i, r_j)$$

More complex metric criteria may be considered if more than one intrinsic parameter is associated with each particle (e.g. ellipsoidal shapes).

Topological Neighborhoods: Voronoi Diagram

In other models, distance-based criteria are not sufficient. This is the case when a spheroidal approximation is not appropriate or a particle represents a physical domain of space that cannot be neglected compared to the size of the particle's local interaction field. Denoting ϕ the ratio of the physical volume of particle i over the volume of its local interaction field, we can state that (a) for planets and their gravitation field, or atoms and their electrostatic potential, ϕ tends to 0; while (b) for objects that interact via their contact surface, as is the case in multicellular systems composing living tissue, then ϕ is much closer to 1 (Fig. 3.8).

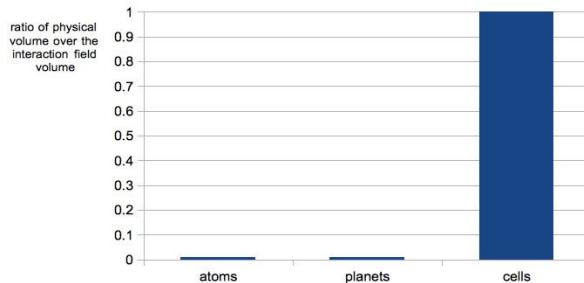


Figure 3.8: Ratio of the physical volume of a particle over the volume of its local interaction field. Atoms and planet have an interaction range that is much larger than their physical envelope whereas cells can only mechanically interact by physically touching their neighbors.

In high- ϕ systems, different criteria must define more specifically what part of the space is occupied by the object associated with the particle. For systems where objects are densely packed against each other, a *Voronoi diagram* of the swarm, $V(P)$, is a convenient solution to determine a decomposition of the space into convex polyhedra $V(i)$ (called Voronoi cells), one per particle (Fig. 3.9):

$$V(i) = \{\vec{X} \in \mathbb{R}^3 : \forall j \in P, j \neq i, \|\vec{X} - \vec{X}_i\| \leq \|\vec{X} - \vec{X}_j\|\}$$

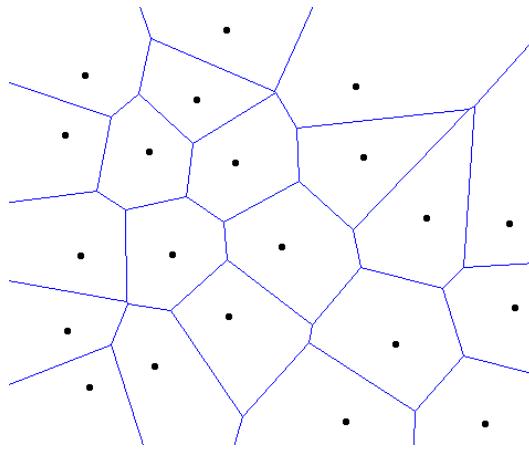


Figure 3.9: 2D Voronoi diagram generated with the Voronoi sweeping algorithm applet by H. Liefke.

In \mathbb{R}^3 , a Voronoi cell is delimited by Voronoi facets that contain all points equidistant from two neighboring cells. A Voronoi edge is shared by three Voronoi cells and a Voronoi vertex is shared by four Voronoi cells. Certain “degenerate” configurations of particles in a swarm perturbs a proper Voronoi diagram construction, for example: three colinear particles, four particles on a circle, or five particles on a common sphere. In inherently stochastic systems with real-valued coordinates, however, these conjunctions have a next to nil probability to occur. Finally, a few Voronoi cells at the periphery of the swarm have an open volume in \mathbb{R}^3 . They share open facets and edges among each other and form the outer shape of the swarm called “convex hull”.

Dual of the Voronoi Diagram: Delaunay Triangulation

From the Voronoi diagram, a fundamental neighborhood rule in computational geometry can be deduced: the *Delaunay triangulation*. When the elements of the Voronoi diagram are polyhedra, the edges that link each center of mass of adjacent polyhedra constitute the dual of the Voronoi diagram—and vice-versa. One can be deduced deterministically from the other (Fig. 3.10).

The process of cutting out space with geometric shapes leaving no overlaps or holes is called *tessellation*. Most of the time, cells remain in contact, each polyhedron’s face is shared by two neighboring cells. Sometimes, however, holes may appear or disappear as cells move during embryogenesis. These holes can be construed as empty polyhedra, therefore we will continue using the term tessellation

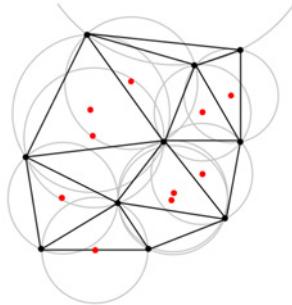


Figure 3.10: A Delaunay triangulation in the 2D plane.

Black dots are the particles, black lines their Delaunay neighborhood links, and red dots the circumcircles' centers. No circumcircle contains a particle in its interior. Image from http://en.wikipedia.org/wiki/Delaunay_triangulation

anyway.

Attraction/Repulsion Potential

All particle-based models' interaction potentials share common features: repulsion at short distance, until a certain equilibrium distance r^{eq} , and attraction beyond, until the border of the interaction field, whether it is finite or not.

Equilibrium Distance

Two different characterization of equilibrium distances can co-exist: (a) equilibrium distances can be implicitly obtained by the relaxed state of the dynamical system, which is a function of its physical properties (elasticity, adhesion, etc.), (b) they can be explicitly specified by geometric calculus.

Implicit Equilibrium Distance

As soon as two cells touch each other, adhesive forces at their surface tend to increase the contact area until inner resistance compensate them completely. This implicitly defines an equilibrium surface. The most influential field of research for describing this phenomenon is called “contact mechanics”. Initiated by the articles of Heinrich R. Hertz entitled “On the contact of elastic solids” (1881) and “On the Contact of Rigid Elastic Solids and on Hardness” (1882), contact mechanics successfully characterized the deformation of two elastic bodies that touch each other as a function of the load applied to them and their mechanical properties. Originally, Hertz only described the non-adhesive surface case but two concurrent models, Johnson-Kendall-Roberts (JKR) [229] and Derjaguin-Muller-Toporov (DMT) [230],

later added adhesion-related terms to the equation, each one performing better on different materials.

For example, if two elastic spheres with elastic moduli E_i and E_j , Poisson's ratios ν_i and ν_j and radii R_i and R_j , their contact is a circle of radius a . The JKR model establishes a relation between load F_{JKR} and contact radius by:

$$F_{\text{JKR}}(a) = \frac{4E^*a^3}{3R^*} - 2\sqrt{2\pi E^* w_{\text{adh}} a^3}$$

where w_{adh} is the adhesive energy per unit area, and

$$E^* = \left(\frac{1 - \nu_i^2}{E_i} + \frac{1 - \nu_j^2}{E_j} \right)^{-1}, \quad R^* = \left(\frac{1}{R_i} + \frac{1}{R_j} \right)^{-1}$$

The equilibrium surface radius a^{eq} is obtained for $F_{\text{JKR}} = 0$:

$$a^{\text{eq}} = \left(\frac{9\pi w_{\text{adh}} R^{*2}}{2E^*} \right)^{1/3}$$

As the shape of the deformation of the surface is assumed to be parabolic, the equilibrium distance between the two spheroid particle centers can also be inferred (Fig. 3.11).

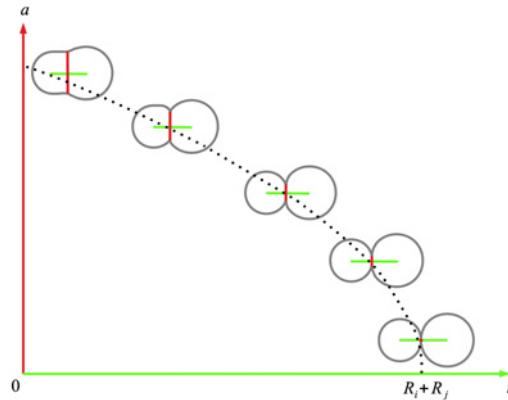


Figure 3.11: Typical profile of the contact radius as a function of the particles' distance in contact mechanics (with adhesion). In abscissa: distance between the particles; in ordinate: contact radius. As the adhesive energy increases or the elastic moduli decrease, the equilibrium contact surface area tends toward the upper left part of the curve.

These equations have already been used in multicellular modeling, but with additional hypotheses that are debatable (e.g. the original Hertz model and the JKR theory both assume that the deformation is small with respect to the radii

of the spheres). Regarding our goal of modeling embryogenesis, a major drawback of contact mechanics is also that it was designed for two interacting spheroidal particles. For multiple particles touching each other, like cells, a contact surface area is surrounded by other contact surface areas, which greatly reduces the free surface.

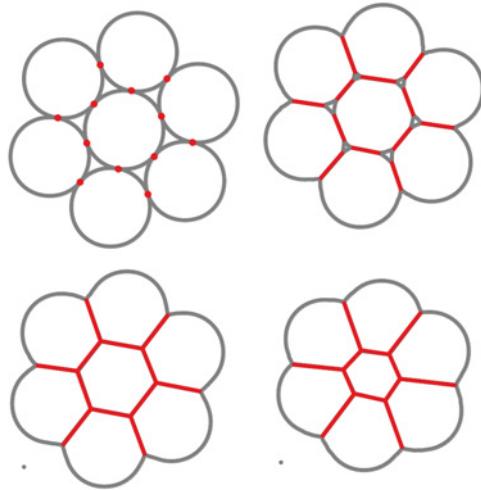


Figure 3.12: Multicellular limits of contact mechanics. As the adhesion energy increases, the contact surface areas become larger and the volume of the central cell is not conserved (bottom right). Contact mechanics does not deal well with multiple surfaces of contact when “triple junction points” appear (from top right to bottom).

A particular problem is that cellular volume tends to decrease when interfacial tension increases (Fig. 3.12). This unfortunately defeats the refined characterization of the surface area equilibrium by models of contact mechanics. One possibility to counterbalance this tendency is to add a volume conservation term, but the outcome is also a serious perturbation of the contact mechanics framework.

In sum, the trouble with the particle-based framework is that the main variables are the coordinates \vec{X}_i of the centers of the particles, and there is no explicit usage of surface areas, for example in the Voronoi calculation. The pair-based framework of neighboring particles does not allow to design more adapted models. Instead of making a pair of particle the generic case and a tuple of particles the exception, we need to do the opposite.

3.2 Hypotheses

This section explains how we express and calculate the mechanical interactions and behavioral properties of the cells. It presents a *discrete-element model using one particle per cell*, driven by an overdamped equation of motion that can be summarily written $\lambda \vec{v}_i = \vec{F}_i^P + \vec{F}_i^{A,\text{int}} + \vec{F}_i^{A,\text{ext}}$ (Section 3.2.1), where \vec{v}_i is the velocity of one cell i , \vec{F}_i^P represents “passive” interaction forces controlling cell stiffness and adhesion, and $\vec{F}_i^{A,\text{int}}$ and $\vec{F}_i^{A,\text{ext}}$ are the “intrinsic” and “extrinsic” components of “active” interaction forces such as protrusive activity or apical constriction. These forces are calculated by summing over a neighborhood \mathcal{N}_i containing nearby cells that are in contact with cell i . This neighborhood is itself defined by metric and topological criteria (Section 3.2.2). \vec{F}_i^P is a *relaxation* force derived from an attractive/repulsive, elastic-like interaction potential (Section 3.2.3), while $\vec{F}_i^{A,\text{int}}$ and $\vec{F}_i^{A,\text{ext}}$ compose a *behavioral* force moving the swarm configuration away from equilibrium (Section 3.2.4).

3.2.1 Generic Cell-Centered Equation of Motion

The *equation of motion* is the core of every mechanical model. It relates the spatial variables of the model and time through hypothesized laws that reflect our understanding of the properties of the studied phenomenon. In classical mechanics, which describes the world of rigid bodies at a macroscopic scale, Newton’s three laws of motion (inertia, acceleration, and action/reaction) constitute the fundamental principles upon which equations of motion are built. In a biomechanical system such as multicellular embryogenesis, however, a modification of these laws is necessary to adapt them to different objects and environments: soft bodies and high viscosity at a near-molecular scale. This section explains our proposal of simplified equations of local cell motion, derived from classical mechanics, in the perspective of applying them to the agent-based modeling and computational simulation of biological development.

Newton’s second law states that the acceleration \vec{a} of a body is proportional to the sum of the forces \vec{F} applied to it. The proportionality factor is the inverse of the mass of the body, a quantity assumed to remain constant over time: $\vec{a} = \vec{F}/m$. In particle-based physics, for each particle characterized by a mass m , acceleration \vec{a} , location \vec{X} , velocity \vec{v} , and radius R , forces are a function of the last three quantities, and Newton’s law reads

$$m\vec{a} = \vec{F}(\vec{X}, \vec{v}, R).$$

At the scale of the biochemical world, however, this classical formulation does not reflect the emergent and rather counterintuitive phenomena that cells exhibit individually and collectively. In fact, cells are so small and their interactions so “sticky” that the physics of their motion is radically different from the one to which

we are accustomed at our scale. The main difference resides in the disappearance of inertial forces. In his article “Life at low Reynolds number”[231], Purcell explains how a cell immersed in water stops moving as soon as he stops pushing it. The cell, an E. Coli bacteria, is so tiny that the viscous forces exerted by the surrounding water completely overwhelm the inertial forces due to motion.

More precisely, when an object of characteristic length R moves at a velocity $v = \|\vec{v}\|$ in a fluid of density ρ and dynamic viscosity μ , the quantity that measures the ratio of the inertial forces to the viscous forces is called the Reynolds number and denoted by Re . Its expression is

$$\text{Re} = \frac{\rho v R}{\mu}$$

where ratio μ/ρ is the kinematic viscosity, denoted by ν , and has a value of $10^{-6} m^2 s^{-1}$ for water. For a person of characteristic length $1m$ swimming at $1ms^{-1}$, Re is of the order of 10^6 . For a bacterium of characteristic length $10^{-6}m$ swimming at $10^{-5}ms^{-1}$, Re is around 10^{-5} . Therefore, the physics of inertia and viscosity is completely different for these two examples. The balance shifts from inertial forces on a person to the viscosity forces on a cell.

A more striking example is the coasting distance of the same bacterium when it stops self-propulsing in water. When only inertial and viscous forces are at work, Newton’s second law becomes

$$m \frac{dv}{dt} = -\gamma_{\text{drag}} v$$

and its solution is [232]

$$v(t) = v_0 e^{-\frac{\gamma_{\text{drag}}}{m} t},$$

which, integrated from $t = 0$ to $t = \infty$ yields the coasting distance d of the bacterium:

$$d = \frac{v_0 m}{\gamma_{\text{drag}}}.$$

For a bacterium swimming at $10^{-5}ms^{-1}$, this distance is 0.1 angstrom. Thus a simple calculation confirms that if the bacterium is not moving by itself or pushed by an external actor, it “immediately” stops moving. Purcell’s interpretation of these results is that If you are at very low Reynolds number, what you are doing at the moment is entirely determined by the forces that are exerted on you *at the moment*, and by nothing in the past.

The consequence is that in a low Reynolds number environment, applied forces produce a velocity, not an acceleration. The displacement is proportional to the instantaneous force and inversely proportional to the a damping coefficient λ . Therefore, these results allow to bypass Newton’s Second Law and we use the following modified equation of motion in our model:

$$\lambda \vec{v} = \vec{F}$$

The above discussion assumed a solid object moving in a fluid flow. In a multicellular environment such as a developing organism, we assume that the same principles apply for the following reasons:

- the dimension of cells are roughly similar to a bacterium (from 300 microns for the zygote to 0.3 micron at the 1024-cell stage, and 0.02 micron when the embryo contains about 15,000 cells)
- the density is equivalent
- the kinematic viscosity of the surrounding cells taken as a flow can be considered even greater than water; for example, at 18°C, the kinematic viscosity of ketchup is 50,000 times higher than water (see [219], Appendix 2 for a similar justification)

Generalizing to a multicellular system, the motion of each cell indexed by i is governed by its interactions with all other cells j belonging to its neighborhood denoted by \mathcal{N}_i , according to

$$\lambda_i \vec{v}_i = \sum_{j \in \mathcal{N}_i} \vec{F}_{ij}$$

A crucial difference with a solid object, however, is that surrounding cells are responsible not only for damping the motion but also for the motion itself. There is no “hand” pushing the cell, as in Purcell’s experiment: the pushing is exerted the cells’ intrinsic behavior and, to progress, it needs to cling and push back surrounding cells a bit like a swimmer pushes back water to move forward. Accordingly, the interaction force \vec{F}_{ij} exerted by cell j over cell i is the sum of three components:

- a *passive* interaction force \vec{F}_{ij}^P , which maintains the integrity of the cell’s volume and controls both the stiffness and the adhesion of the interaction; this force always applies irrespective of the state of the cell but can be modulated via the stiffness and adhesion components (explained in Section 3.2.3)
- an *intrinsic active* interaction force $\vec{F}_{ij}^{A,int}$, which gives rise to specific behaviors such as protrusive activity or apical constriction; this force depends on the state of cell i , in particular vanishes when the cell is “resting”, and
- an *extrinsic active* interaction force $\vec{F}_{ij}^{A,ext}$ induced by the intrinsic force $\vec{F}_{ji}^{A,int}$ of neighbor cell j , i.e. both forces are equal but have opposite directions: $\vec{F}_{ij}^{A,int} = -\vec{F}_{ji}^{A,ext}$ (both explained in Section 3.2.4).

Moreover, the damping coefficient λ_i , which plays here a role somewhat equivalent to the mass m_i in Newton's Second Law, is proportional to the surface of the cell: $\lambda_i = \lambda_0 R_i^2$. Neglecting the cell's weight (proportional to m_i), the final equation then reads

$$\vec{v}_i = \frac{1}{\lambda_0 R_i^2} \sum_{j \in \mathcal{N}_i} \left(\vec{F}_{ij}^P + \vec{F}_{ij}^{A,\text{int}} + \vec{F}_{ij}^{A,\text{ext}} \right)$$

This will be the equation used throughout our study. At times, we will also used the following condensed notations:

$$\begin{aligned} \vec{v}_i &= \frac{1}{\lambda_0 R_i^2} \sum_{j \in \mathcal{N}_i} \left(\vec{F}_{ij}^P + \vec{F}_{ij}^A \right) \quad \text{with } \vec{F}_i^A = \vec{F}_{ij}^{A,\text{int}} + \vec{F}_{ij}^{A,\text{ext}}, \\ \text{or } \vec{v}_i &= \frac{1}{\lambda_0 R_i^2} \left(\vec{F}_i^P + \vec{F}_i^{A,\text{int}} + \vec{F}_i^{A,\text{ext}} \right) \quad \text{with } \vec{F}_i^X = \sum_{j \in \mathcal{N}_i} \vec{F}_{ij}^X. \end{aligned}$$

The following sections will describe in more detail how the neighbors \mathcal{N}_i of each cell i are identified (Section 3.2.2), and then how the forces coming from neighboring cells j will be calculated.

3.2.2 Cell Neighborhood

A cell sustains forces exerted by the cells in contact with it. While the notion of “contact” is straightforward in certain paradigms (such as cellular Potts, see Section 3.1.1), in a model as abstract as the particle-based model that we rely on, a custom *neighborhood algorithm* must be designed. At any time step, the only information that we have about the spatial structure of swarm S is that each one of its cells, represented by a spheroidal particle i , is characterized by a location $\vec{X}_i = (x_i, y_i, z_i)$ and a volume V_i (Fig. 3.13, left). The objective of this section is to infer from these intrinsic parameters the *neighborhood links* and *contact surface areas* between cells (resp. the black and red edges in Fig. 3.13, right). To this aim, we present a two-step strategy: first, a preselection of potential neighbors that obey certain metric criteria, then a refinement of this list according to topological features.

The regular Delaunay triangulation would be in many ways the most adapted method, as it establishes a list of neighbors according to topological criteria that can also be weighted by each particle's size. However, it has the major drawback of being very computationally expensive. Since we rely on a GPGPU architecture (general-purpose graphics processing units, see Annex A) for our simulations and, to our knowledge, there exists no parallel implementation of regular triangulation yet, we have designed a custom algorithmic flow instead. The goal of our neighborhood algorithm is to produce the most topologically plausible structure (a comparison with the regular triangulation is provided at the end of this section). Our method iteratively builds two lists of neighbor cells: the first list \mathcal{N}_i^m is based on metric criteria and then it is filtered by topological criteria to produce the second list \mathcal{N}_i^t .

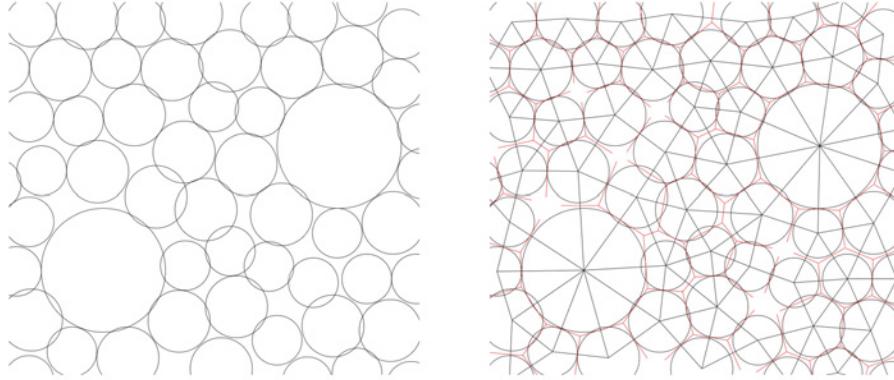


Figure 3.13: Output of the neighborhood algorithm. Left: A 2D swarm of cells characterized by various given positions and radii. Right: The neighborhood links (black edges) and contact surface areas (red edges) between cells calculated by the method described in this section.

First Approximation: Metric Neighborhood

Two cells of radius R_i , R_j and locations \vec{X}_i , \vec{X}_j are considered to be neighbors according to a metric criterion if and only if their relative distance is less than the sum of their radii $R_i + R_j$ multiplied by a constant factor c_{\max} . Thus the metric neighborhood set of cell i is defined as follows:

$$\mathcal{N}_i^m = \left\{ j \in S : \left\| \vec{X}_i - \vec{X}_j \right\| \leq c_{\max}(R_i + R_j) \right\}$$

We denote the distance between the two cells by $r_{ij} = \left\| \vec{X}_i - \vec{X}_j \right\|$ and the cutoff value $r_{ij}^{\max} = c_{\max}(R_i + R_j)$ so that, equivalently, j is a metric neighbor of i if and only if $r_{ij} \leq r_{ij}^{\max}$. Although cells are considered “spheroidal” here, we add the possibility for them to deform in order to reach farther away neighbor cells. Radius r_{ij}^{\max} sets the maximum distance of this deformation, i.e. the spherical domain of space inside which cells can still potentially interact (Fig. 3.14).

Our next step will be to evaluate c_{\max} . For this, we establish an empirical law relating the distance r_{ij} between two cells i and j to their contact surface area, denoted by A_{ij} , such that at the cutoff point $r_{ij} = r_{ij}^{\max}$, A_{ij} is zero.

Empirical Estimation of the Contact Surface Areas

In the absence of real biological data, we construct here an artificial testbed experiment based on the weighted Delaunay triangulation to infer an approximate relationship between cellular distances and contact surface areas. We will then generalize and apply this law to our multicellular model of embryogenesis. In the present experiment, an ellipsoidal domain of space is filled with three consecutive

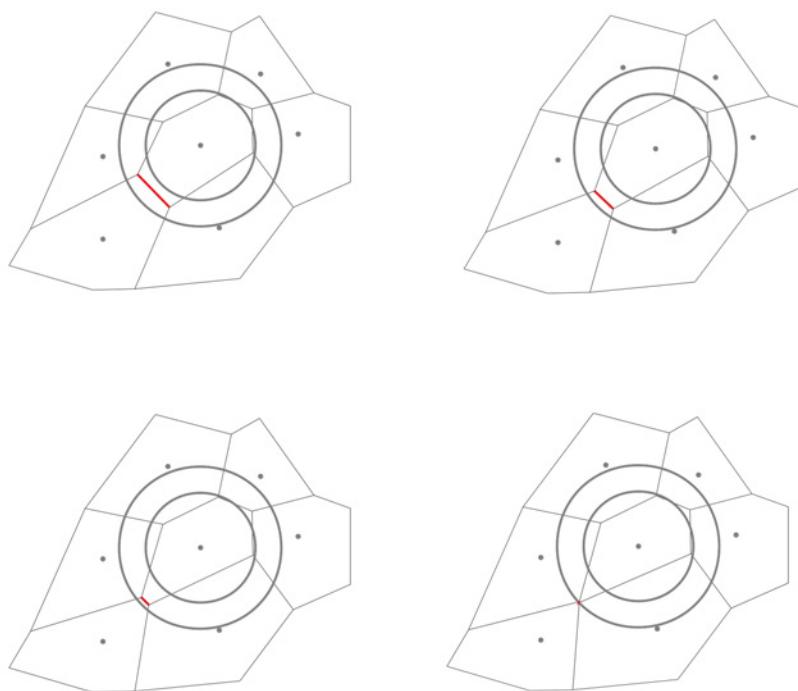


Figure 3.14: Schema of the particle-cell and its maximum distance of deformation. The smaller circle gives the radius of the spherical cell and the larger circle gives the maximum distance of deformation.

generations of cell division g_1, g_2, g_3 in equal quantities. Their volumes are respectively V_{g_1} , $V_{g_2} = (0.5)V_{g_1}$, and $V_{g_3} = (0.5)^2 V_{g_1}$. All cells are considered spheroidal, and their average radii are R_{g_1} , $R_{g_2} = (0.5)^{1/3} R_{g_1}$, $R_{g_3} = (0.5)^{2/3} R_{g_1}$. These radii are the coefficients used in the weighted Delaunay triangulation. The number N of cells per generation is adjusted so that the volume of the ellipsoid container is fully occupied by the $3N$ cells packed together. Then the ellipsoid is deformed, forcing the cells into various spatial rearrangements (Fig. 3.15 and Movie 3.3). Deformation happens by making the first semi-principal axis of the ellipsoid oscillate between two extreme values in a sine motion, while the second axis remains constant and the third axis is adjusted so that the volume of the ellipsoid remains constant.

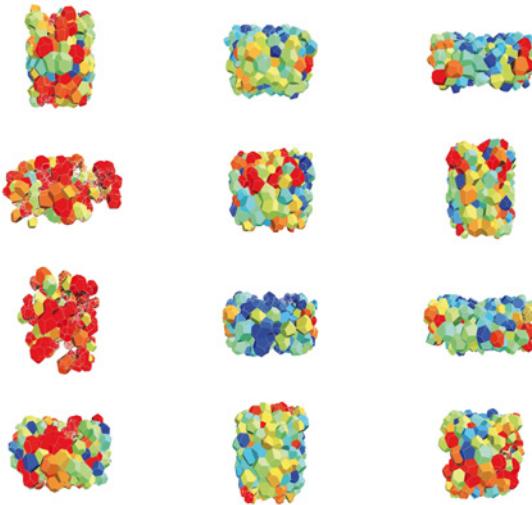


Figure 3.15: Estimation of the contact surface area in a weighted Delaunay/Voronoi simulation. An ellipsoid filled with 300 cells is artificially deformed (stretched and compressed) to generate various distance-surface measurement pairs. Blue cells indicate local compression (i.e. weighted Voronoi volume less than theoretical volume), red cells indicate local dilation.

Fig. 3.15 shows the Voronoi volumes in the interior of the ellipsoid only, i.e. the ones not in contact with the convex hull of the cell pack. The color code represents the normalized difference between the theoretical volume of the cell (specified by its radius) and the observed volume of the cell among its neighbors (specified by weighted Delaunay/Voronoi): $c = V_{\text{obs}}/V_{\text{theo}} - 1$. The lower and more negative the difference (i.e. compressed neighborhood), the closer to blue; the higher and more positive (i.e. dilated neighborhood), the closer to red. In the middle of the distribution, greenish cells correspond to a small $|c|$ ratio (less than 5).

We index by g an experiment that uses a set of $3N$ cells from generations $g_1 = g$, $g_2 = g + 1$ and $g_3 = g + 2$. Each experiment g produces 6 sets of pairs

$\{(r_{ij}, A_{ij})\}_{(i,j) \in g' \times g''}$ depending on the types (g', g'') of adjacent cells (i, j) , where $(g', g'') = (g_1, g_1), (g_1, g_2), (g_1, g_3), (g_2, g_2), (g_2, g_3)$, or (g_3, g_3) . We carried out 6 experiments corresponding to $g = 9, 10, \dots, 14$, representing roughly the zebrafish division stages in which three consecutive generations can effectively be mixed. Thus we produced a total of 36 experimental plots, each one containing about 10,000 points, in the $\{(r, A)\}$ plane. Fig. 3.16 shows one of these plots when $(g, g', g'') = (9, 10, 11)$.

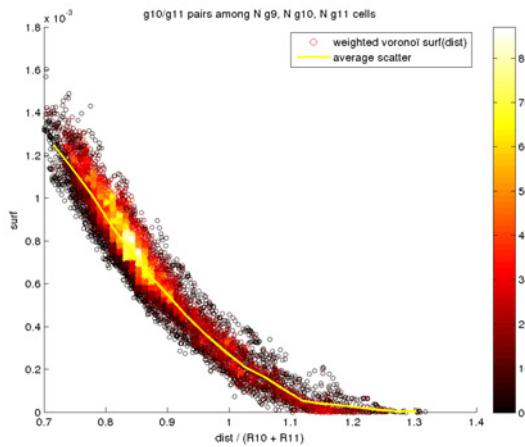


Figure 3.16: Scatter plot of the contact surface area between neighboring Voronoi domains as a function of the distance between their centers. Each dot is a surface-distance pair of values obtained from two neighbors according to the weighted Voronoi algorithm. The color map reflects the density in plot space (white for highest density). The central yellow curve is the regression of the distribution (segmented curve calculated at discrete distance intervals).

A first observation is that, as expected, most distances r_{ij} are lower than $R_i + R_j$ (i.e. points are to the left of a vertical line passing through 1). Now we use a parabola for the shape of the empirical curve that will be fitted to this data:

$$A(r_{ij}, R_i, R_j) = \begin{cases} a (r_{ij} - r_{ij}^{\max})^2 & \text{if } r_{ij} < r_{ij}^{\max} \\ 0 & \text{if } r_{ij} \geq r_{ij}^{\max} \end{cases}, \quad \text{where } r_{ij}^{\max} = c_{\max}(R_i + R_j)$$

We decided to use a monomial of degree 2 to make it scale like a surface. If the distance between two double-size neighboring cells is doubled, the area of the surface of contact will quadruple.

$$A(2r_{ij}, 2R_i, 2R_j) = 4A(r_{ij}, R_i, R_j)$$

Finally, to distill the optimal pair of parameters (a, c_{\max}) from these data plots, we measure the discrepancy between the generated data and the empirical law based on the sum of squared errors of prediction (SSE), denoted by E :

$$E(a, c_{\max}) = \sum_{i,j} (A_{ij} - A(r_{ij}, R_i, R_j))^2$$

In our case where the sample of variable is split into 6 experiments and 6 pairs of cell types per experiment, the SSE is calculated as follows:

$$E(a, c_{\max}) = \frac{1}{6} \sum_{g=9}^{14} \left(\frac{1}{12} \sum_{(g', g'')=(g,g)}^{(g+2, g+2)} \left(\frac{1}{n(g', g'')} \sum_{(i,j) \in g' \times g''} (A_{ij} - a(r_{ij}, c_{\max}(R_{g'} + R_{g''}))^2) \right) \right)$$

We performed an exhaustive exploration of all the pairs of parameters (a, c_{\max}) until the 4th decimal to look for the lowest E value. A heatmap of this search can be visualized in Fig. 3.17.

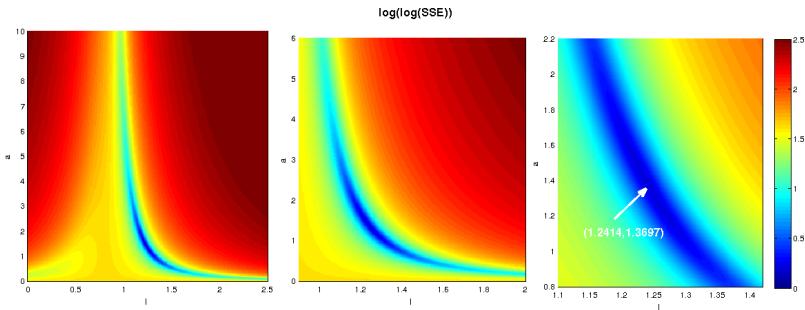


Figure 3.17: Heat map of the sum of squared errors of prediction (SSE) for various pairs of parameters (a, c_{\max}) . Blue regions indicate the best fitness domains, with point $(a = 1.3697, c_{\max} = 1.2414)$ representing the absolute best fit. In abscissa: c_{\max} ; in ordinate: a ; in color: $\log(\log(\text{SSE}))$ to accentuate the contrast.

In conclusion, we found that the best pair of values was $(a = 1.3697, c_{\max} = 1.2414)$ and will use them throughout our simulations. A comparison of this empirical law with the generated data is shown in 3.18 on a selection of 6 different plots out of the 36.

Topological Refinement of the Metric Neighborhood

Now that we have estimated a numerical value for c_{\max} , we can calculate each metric neighborhood \mathcal{N}_i^m according to its definition above (coefficient a will be used in the next section when we need to calculate the forces based on contact surface areas).

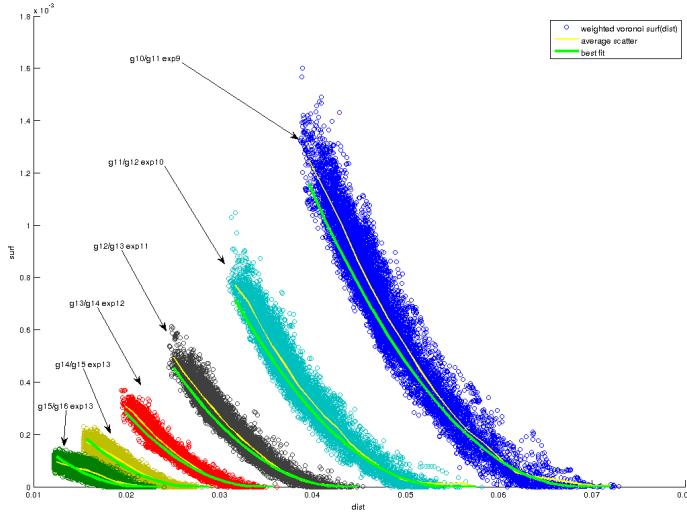


Figure 3.18: Comparison of the best fit with measures from various experiments.

As described in the state of the art, however, a purely metric neighborhood is not viable as it often leads to volumes collapsing during simulation when the adhesion between interacting cells is high. This is why in this section we are going to filter \mathcal{N}_i^m through topological criteria to obtain \mathcal{N}_i^t .

The 3D Delaunay triangulation could do the job but is much too slow to compute because it involves heavy data structures composed of multiple tetraedra. Our own solution is inspired by a similar neighborhood model: the *Gabriel graph*. In 2D, whereas the Delaunay triangulation imposes that no node be found inside the circumcircle of any triangle, the Gabriel graph method imposes that no node be found inside the circle whose diameter is a valid neighborhood edge (Fig. 3.19).

Here, we propose to generalize the Gabriel criterion to the 3D case straightforwardly: two cells (i, j) will be considered neighbors if no other cell k of the swarm is located inside the sphere whose diameter is the segment \overline{ij} , i.e. with origin $(\vec{X}_i + \vec{X}_j)/2$ and diameter $r_{ij} = \|\vec{X}_i - \vec{X}_j\|$:

$$\mathcal{N}_i^t = \left\{ j \in \mathcal{N}_i^m : \forall k \in \mathcal{N}_i^m, \left\| \vec{X}_k - \frac{1}{2}(\vec{X}_i + \vec{X}_j) \right\| \geq \frac{1}{2} \left\| \vec{X}_i - \vec{X}_j \right\| \right\}.$$

To our knowledge, the term Gabriel graph has not been formally defined in 3D. However, when it happens to be mentioned [233], triangles are used instead of segments to define circumspheres. This makes sense, as in n dimensions, the Delaunay triangulation uses n -simplices for its basic units (triangles in 2D, tetrahedra in 3D) while the Gabriel graph uses $(n - 1)$ -simplices (edges in 2D, triangles in 3D). Here, however, we will continue using edges in 3D while still referring to our method as

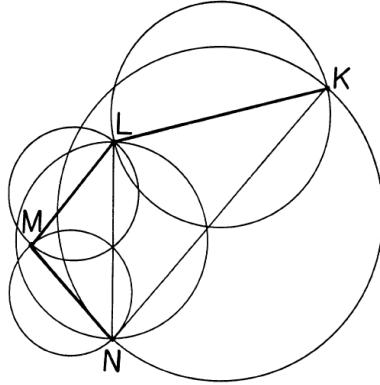


Figure 3.19: An example of Gabriel graph on four nodes K, L, M and N. Only edges KL, LM and MN are valid because the disks whose diameters are these edges do not contain any other node (whereas, for example, the disk based on LN contains node M).

a Gabriel graph. Note that in the 2D case, the Gabriel graph is a subgraph of the Delaunay triangulation. Every pair of neighbors in a Gabriel graph is also a pair of neighbors according to the Delaunay criteria (Fig. 3.20).

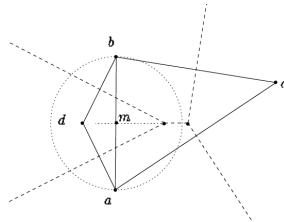


Figure 3.20: A 2D Gabriel graph is a subset of a Delaunay triangulation. Here, 5 solid lines represent the Delaunay edges. The Gabriel graph has only 4 of them because it excludes edge ab . Visually, a Delaunay edge is also a Gabriel edge if it intersects the contact border (dashed lines) between the two cell domains (schema from [234])

3.2.3 Interaction Potential

As explained in Section 3.2.1 above, two neighboring spheroidal cells i, j of radii R_i and R_j separated by a distance r_{ij} interact in our model through forces \vec{F}_{ij} (exerted by j over i), which contain a passive *relaxation* term \vec{F}_{ij}^P . The latter is derived from an attractive/repulsive interaction potential, leading the swarm of cells toward an equilibrium state in the absence of any other forces. As it will be explained below, the underlying rationale of this potential is to *maximize the contact surface area*

between every neighboring cells in the swarm.

For every pair of neighboring particles (i, j) we focus in this section on the passive interaction potential E_{ij}^P from which \vec{F}_{ij}^P is derived, and model it from three juxtaposed domains:

- a *repulsion* domain (decreasing E) at distances shorter than a certain equilibrium distance defined by $r_{ij}^{\text{eq}} = c_i^{\text{eq}}R_i + c_j^{\text{eq}}R_j$, where in the most general case c_i^{eq} is a coefficient that can vary from cell to cell
- an *attraction* domain (increasing E) at distances greater than r_{ij}^{eq}
- a *neutral* domain (constant E) beyond the maximum limit of the interaction field $r_{ij}^{\text{max}} = c_{\text{max}}(R_i + R_j)$

In the previous section, we established the value of the maximum distance r_{ij}^{max} , by estimating an average universal value for coefficient c_{max} based on empirical statistics of the contact surface areas between neighboring particles. Equipped with this rule, we then performed a two-step strategy of selection of the neighboring cells, which gave us the final topological-metric neighborhood of i , \mathcal{N}_i^t .

Here, we want to calculate the relaxation forces \vec{F}_{ij}^P , which are based on an estimation of the equilibrium coefficients c_i^{eq} . As reviewed in Section 3.1, the definition of the equilibrium distance can be stated either explicitly or implicitly. In the implicit case, for example in the most advanced models of contact mechanics, it results from a trade-off between different biomechanical properties. We showed, however, that these models ran against major difficulties when multicellular interactions are considered. Instead, we decide here to explicitly identify the distance of equilibrium between neighboring cells with the *distance between their locations in the densest arrangement of sphere packing*. We assume that intercellular material covers no significant volume (otherwise, additional particles representing the extra-cellular material would have to be added, which we do not do here).

In the 2D plane, the densest packing of identical disks is the hexagonal lattice (Fig. 3.21). The ratio of the surface occupied by the disks to the surface containing the disks is equal to the ratio of the surface of a disk to the surface of an hexagon (its Voronoi domain), i.e. $\pi/(2\sqrt{3}) \simeq 0.9068997\dots$

In 3D, Kepler's conjecture states that the hexagonal close packing and face-centered cubic close packing arrangements of equally sized sphere have the greater density of all arrangements (proven by Thomas Hales in 1998). A "hexagonal close packing" arrangement is obtained by superposing and staggering two 2D hexagonal lattices of sphere in an ABAB... pattern (Fig. 3.22).

In this arrangement, each sphere touches 12 other spheres, i.e. each cell has 12 neighbors (six neighbors on the 2D plane and three other neighbors on each of the surrounding planes). The typical tile of the dual space-filling tessellation (Voronoi domains) is called a trapezo-rhombic dodecahedron (TRD; Fig. 3.23).

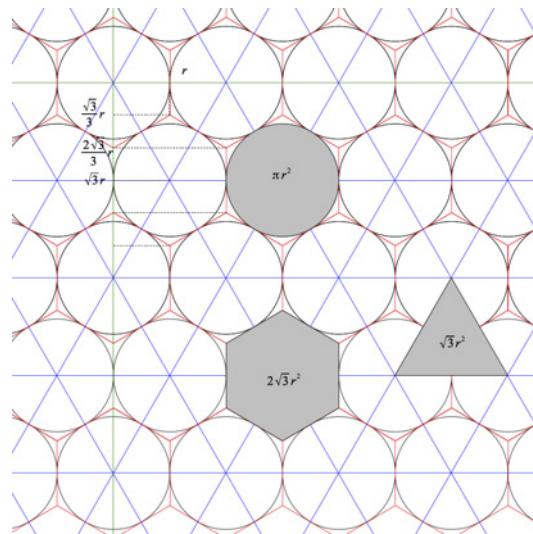


Figure 3.21: Perfect hexagonal lattice. In the 2D plane, the hexagonal lattice configuration realizes the densest packing of disks.

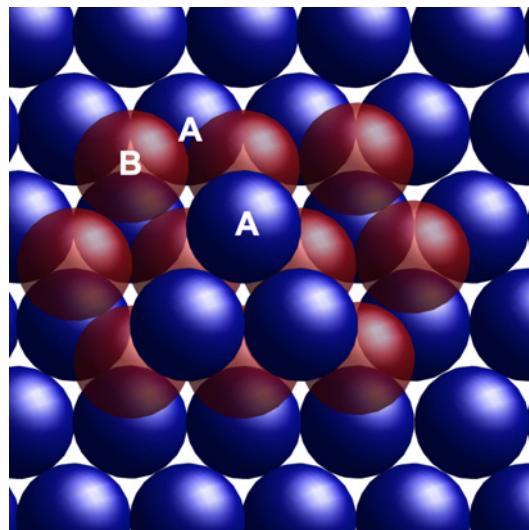


Figure 3.22: Hexagonal close packing in 3D.

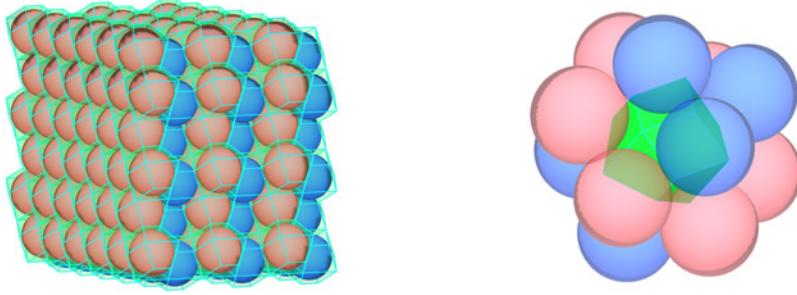


Figure 3.23: Trapezo-rhombic dodecahedra (TRD) constitute the Voronoi domains of a maximally compact 3D packing of equally sized solid spheres.

In this case, the atomic packing factor is the ratio of the volume of a sphere to the volume of its surrounding TRD and is equal to $\pi/(3\sqrt{2}) \simeq 0.74048\dots$. Therefore, in the perfect and uniform arrangement here, the equilibrium distance between any cells (i, j) of equal radii $R_i = R_j = R$ is constant and reads

$$r^{\text{eq}} = 2c_{\text{eq}}R, \quad \text{with } c_{\text{eq}} = \left(\frac{\pi}{3\sqrt{2}}\right)^{1/3} \simeq 0.904699895\dots$$

We now generalize the notion of distance of equilibrium to pairs of unequally sized neighbor cells. In this case, the common radius of equilibrium depends on individual contributions from each cell, which will be simply proportional to their radius, keeping the same constant coefficient c_{eq} above (Fig. 3.24):

$$r_{ij}^{\text{eq}} = c_{\text{eq}}R_i + c_{\text{eq}}R_j = c_{\text{eq}}(R_i + R_j)$$

In the previous sections, we have established that the contribution of a given cell i of radius R_i to the equilibrium distance with neighboring cells was $r_i^{\text{eq}} = c_{\text{eq}}R_i$ and its contribution to the maximal distance was $r_i^{\text{max}} = c_{\text{max}}R_i$, where $c_{\text{eq}} \simeq 0.91$ and $c_{\text{max}} \simeq 1.241$. As illustrated in Fig. 3.25, we can now define the interaction potential E_i^P around a given cell i to be repulsive for a distance below r_i^{eq} (blue area), attractive between r_i^{eq} and r_i^{max} (red area), then constant beyond r_i^{max} .

Given these boundary conditions, we consider here the simplest form of schematic mechanical model that can realize both the short-range repulsion and long-range attraction parts, namely: *spring-like forces derived from an elastic potential*. Thus forces \vec{F}^P will be a linear function, and potentials E^P a quadratic function of $(r - r^{\text{eq}})$.

Mechanical interactions between neighboring cells have been extensively studied by the biophysical scientific community, yet due the multiprotein and polymeric

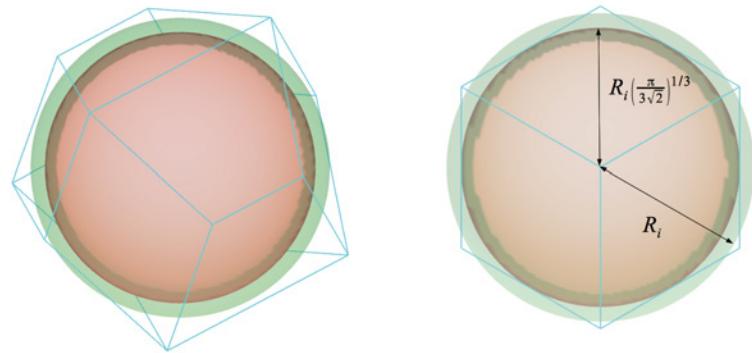


Figure 3.24: The green sphere has the same volume as the trapezoidal rhombic dodecahedron. The orange sphere has a radius $c_{eq}R_i$ tangential to each face of the TRD.

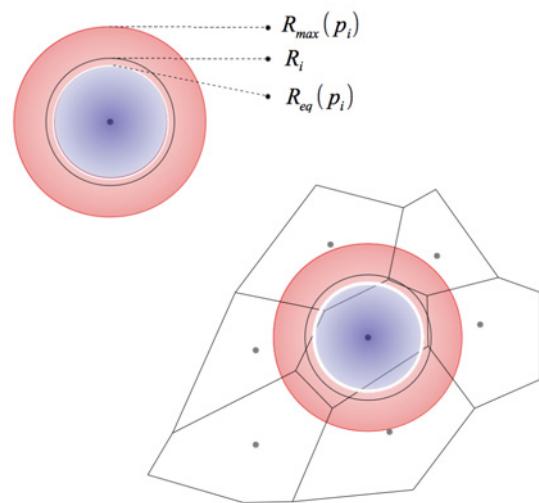


Figure 3.25: Schema of the particle-cell, its equilibrium distance and its maximum distance of deformation.

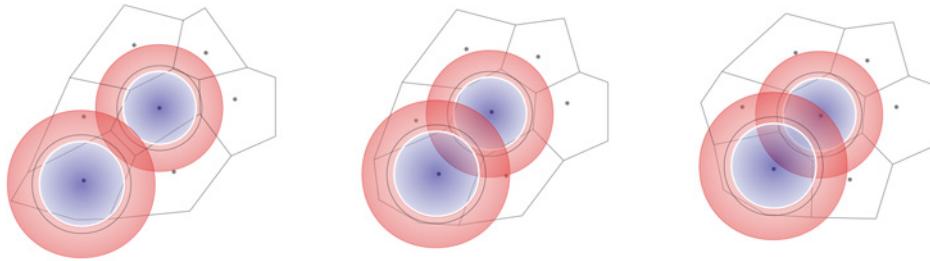


Figure 3.26: Neighbor particle-cells in attraction (left), equilibrium (middle) and repulsion (right) configuration.

nature of the cytoskeleton, membrane flexibility and adhesiveness, it remains an extremely elusive topic difficult to summarize by a definitive set of equations. Roughly, one consensus hypothesis is that a major driving mechanism can be characterized by intercellular *surface tension*. The two main components of this tension are cellular *adhesion*, which tends to lower it, and cellular *cortical tension*, which acts antagonistically. The balance between these two components determines the ultimate behavior of the interaction, i.e. attractive or repulsive.

Another fundamental mechanism must also be taken into account: *volume preservation* in cells. As we made the assumption that cells are spheroidal and cell interactions can be decoupled into two individual, additive contributions, we assume here that the volume conservation principle participates in the *repulsive* part of the potential only. This allows us to introduce a correction in the form of a discontinuity in the force derivative at the distance of equilibrium r_i^{eq} . We will also globally refer as “adhesion” to the cumulated effect of true surface adhesion and its antagonist, cortical tension.

Accordingly, to modulate the intensity of adhesion, we split the expression of F^P into two parts, one below and one above the equilibrium distance r_i^{eq} , corresponding to two different stiffness coefficients w (Fig. 3.27). A low (resp. high) adhesion coefficient w will induce a weak (resp. strong) attraction:

$$\vec{F}_{ij}^{P,\text{lin}} = \begin{cases} -w_{\text{rep}}(r_{ij} - r_{ij}^{\text{eq}}) \cdot \vec{u}_{ij} & \text{if } r_{ij} < r_{ij}^{\text{eq}} \\ -w_{\text{adh}}(r_{ij} - r_{ij}^{\text{eq}}) \cdot \vec{u}_{ij} & \text{if } r_{ij} \geq r_{ij}^{\text{eq}} \end{cases}$$

Cells (i, j) start to adhere only when their relative distance becomes less than r_{ij}^{max} . As the intensity of the attraction can not be suddenly maximal at that point, a half-bell shape between r_{ij}^{eq} and r_{ij}^{max} is better suited to model a more realistic cellular interaction. We obtain this shape by multiplying the linear force $\vec{F}_{ij}^{P,\text{linear}}$ by the area of the surface of contact A_{ij} elaborated in Section 3.2.2 and recalled in Fig. (Fig. 3.28).

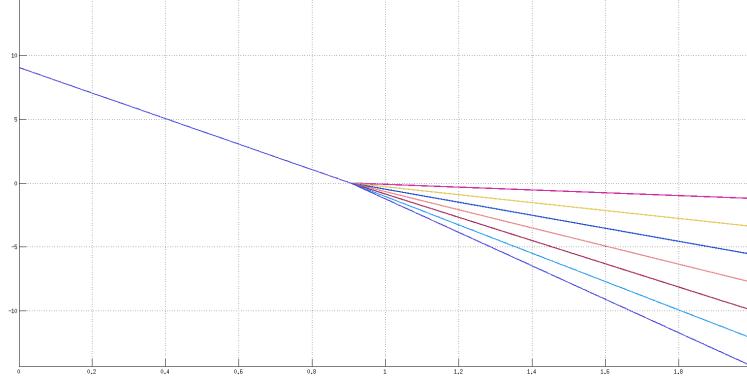


Figure 3.27: Plot of the linear spring-like force $\vec{F}^{P,\text{lin}}$. The repulsion coefficient w_{rep} is fixed, while the adhesion coefficient w_{adh} can vary. Here, the equilibrium distance is $r^{\text{eq}} = 1$.

$$A_{ij} = A(r_{ij}, R_i, R_j) = \begin{cases} a \left(r_{ij} - r_{ij}^{\max} \right)^2 & \text{if } r_{ij} < r_{ij}^{\max} \\ 0 & \text{if } r_{ij} \geq r_{ij}^{\max} \end{cases}, \quad \text{where } r_{ij}^{\max} = c_{\max}(R_i + R_j)$$

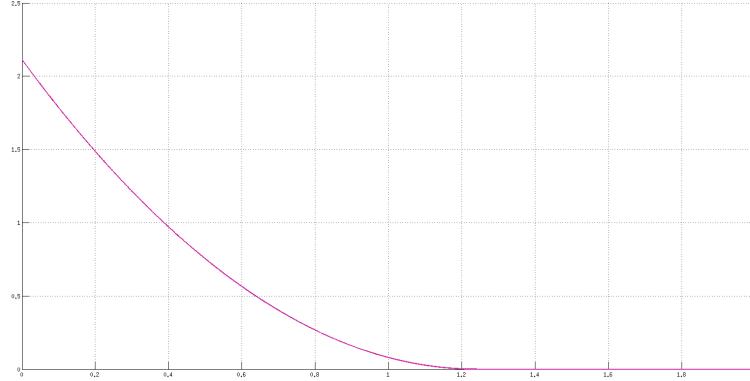


Figure 3.28: Contact surface area A_{ij} as a function of inter-cellular distance r_{ij} . Here, the sum of R_i and R_j is 1.

In the whole $[0, r_{ij}^{\max}]$ interval, the depth of the well is controlled by a repulsion coefficient w_{rep} (until r_{ij}^{eq}) and an adhesion coefficient w_{adh} (after that), so that the final expression of the nonlinear version of the relaxation force that takes into account the contact surface area reads (Fig. 3.28):

$$\vec{F}_{ij}^P = A_{ij} \cdot \vec{F}_{ij}^{P,\text{lin}} = \begin{cases} -aw_{\text{rep}}(r_{ij} - r_{ij}^{\max})^2(r_{ij} - r_{ij}^{\text{eq}}) \cdot \vec{u}_{ij} & \text{if } r_{ij} < r_{ij}^{\text{eq}} \\ -aw_{\text{adh}}(r_{ij} - r_{ij}^{\max})^2(r_{ij} - r_{ij}^{\text{eq}}) \cdot \vec{u}_{ij} & \text{if } r_{ij} \geq r_{ij}^{\text{eq}} \text{ and } r_{ij} < r_{ij}^{\max} \\ \vec{0} & \text{if } r_{ij} \geq r_{ij}^{\max} \end{cases}$$

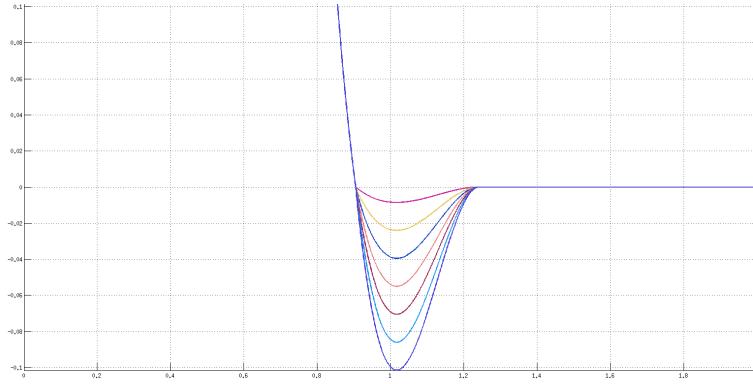


Figure 3.29: Plot of the relaxation force \vec{F}_i^P with variable adhesion coefficient w_{adh} .

Finally, to compare our custom attraction/repulsion force \vec{F}_{ij}^P with classical potentials, we show their superimposed plots in Fig. 3.30.

3.2.4 Cell Specific Behavioral Forces

In addition to the passive, equilibrium-oriented attraction/repulsion forces \vec{F}_{ij}^P , which were calculated in the previous section, we introduce proactive behavioral forces, via an “intrinsic” term $\vec{F}_{ij}^{A,\text{int}}$ and its “extrinsic” counterpart $\vec{F}_{ij}^{A,\text{ext}}$ (explained below), whose purpose is to represent a schematic model of the cells’ specialized biomechanics. During development and across numerous species, cells manifest a wide variety of mechanical properties and behavioral phenomena. While many different taxonomies have been proposed, we will rely in our work on the cell behavior ontology (CBO) described in Section 5.1.1. Moreover, inside this diverse repertoire, we focus here on one particular mechanism that we believe is the main driving force in the zebrafish early gastrulation, namely: *cellular protrusion*.

Our mechanical model offers an abstraction of the cell as the microscopic scale of embryogenesis, based on the principle that *a protrusion is effective only if it has a support to exert its action upon*. (Note that we came up with this idea and designed the force model described below independently from a similar study that was published recently [216].) In order to explain how we interpret the protrusive

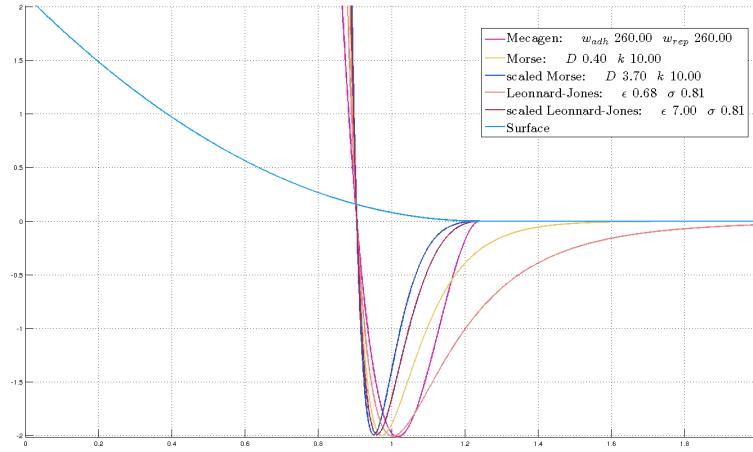


Figure 3.30: Comparison of the relaxation force used in MECAGEN with various classical attraction-repulsion forces. (See legend inset.)

activity of a cell at this schematic scale, however, we need to go down one level and dissect the mechanisms responsible for protrusion at the subcellular level. The following description reflects our understanding of protrusion gathered from many readings on this topic in the biomechanical literature. It is still an object of intense scrutiny and debate among researchers.

In summary, cell protrusion is essentially a *cyclic* activity, somewhat similar to the activity of tracked vehicles (such as tanks) except that, as gravity plays almost no part in cellular interactions, adhesion is regulated in a special way to avoid sliding between the cell surfaces in contact. The protrusive activity induces an *intercalation* of the cell between its neighbors. One requirement for displaying protrusive activity is *polarization*. Cells need to have an asymmetrical distribution of internal substances to specify the direction in which they protrude. How the polarization axis is determined in our model will be explained in Chapter 5 on mechanogenetic coupling. For now, we assume it is a given parameter. This polarization axis determines two regions, or *poles*, of the cell belonging to the extremities. Each pole can be the location of protrusive activity: if only one pole is active, the cell is called “monopolar”; if both are active, it is called “bipolar”.

At the subcellular level, the main structure underlying protrusion is the cellular scaffolding called *cytoskeleton*, whose three main components are “microfilaments” (similar to an envelope that contracts and dilates), “intermediate filaments” (similar to cables exerting a tension) and “microtubules” (similar to beams resisting compression). To explain protrusion, we focus on the microfilaments, which for the most part constitute the “cell cortex”, a mesh-like network essentially made of actin

and myosin molecules that lie just below the plasma membrane and are attached to it by catenin molecules. Actin proteins are rangy molecules that we can consider, metaphorically speaking, the “bones” of the cell cortex, while myosin molecules play the role of “muscles” and “ligaments” linking these “bones”. When combined, actin and myosin create a strong and flexible “tensegrity” network that can resist internal and external pressure. Moreover, actin is a polymer which can disassemble itself at one of its extremities and reassemble at the other. This allows an active deformation of the acto-myosin network which, as illustrated in Fig. 3.31, is the driving mechanism of protrusion.

A cell is attached to its neighbors by molecular bridges (thin orange edges in the illustration), which are essentially cadherin molecules, and actively deform their internal acto-myosin cortical network in the direction of the polarization axis (blue lines and curves). A bulge eventually appears at the active pole, pushing away neighboring cells. No intercalation process would be observed, however, without an accurate regulation of the adhesion contacts between the cells. Thus, in addition to the regular adhesion bonds, special *focal adhesion points* (thick red edges with numbers in the illustration) also appear at the surface of the protruding region of the cell. These bonds bear the extra load generated by the protrusive activity of the acto-myosin cortex. They become visible around the tip of the bulge and, as the cell is advancing, maintain spatial cohesion between neighboring cytoskeletons. Without them, the cell would slip on the surface, or at least the efficiency of the protrusion would be greatly reduced. Focal adhesion points gradually disappear from the cell membrane as the cell advances relatively to the bonded neighbors.

The coupled action of acto-myosin cortex and the focal adhesion points form a sort of “treadmill” originating at the tip of the protrusion (Fig. 3.32). This movement induces a torque transmission between the protruding cell and each of the neighbor cells attached by focal adhesion points. The transfer of cell material in the bulge results from the “intrinsic” force generated by the acto-myosin cortex, while the “extrinsic” force is exerted on the adjacent cell through the focal adhesion points. In some cases, an additional mechanism of cellular contraction at the back of the cell amplifies the intrinsic force (not included here). We assume that the distribution of focal adhesion points is homogeneous on the surface of the cell, so the quantity of torque transmitted between two neighboring cells i and j is proportional to their contact surface area $A_{ij} = A(r_{ij}, R_i, R_j)$.

In our particle-based framework, the mathematical interpretation and representation of this mechanism is the following (**Fig. 3.33A**): a cell i has a normalized *polarization axis* \vec{U}_i that we consider given in this part (black arrow). Later, in Chapter 5 integrating biomechanics and molecular/genetic dynamics, we will explain how this polarization axis can be determined by feedback from molecular signaling, neighboring cell-cell contacts, and/or the active forces themselves (see Section 5.1.4). Among the neighboring cells in \mathcal{N}_i^t , we denote by \mathcal{N}_i^{t+} the sublist of neighbors that make contact with cell i on its “positive” pole, i.e. which are

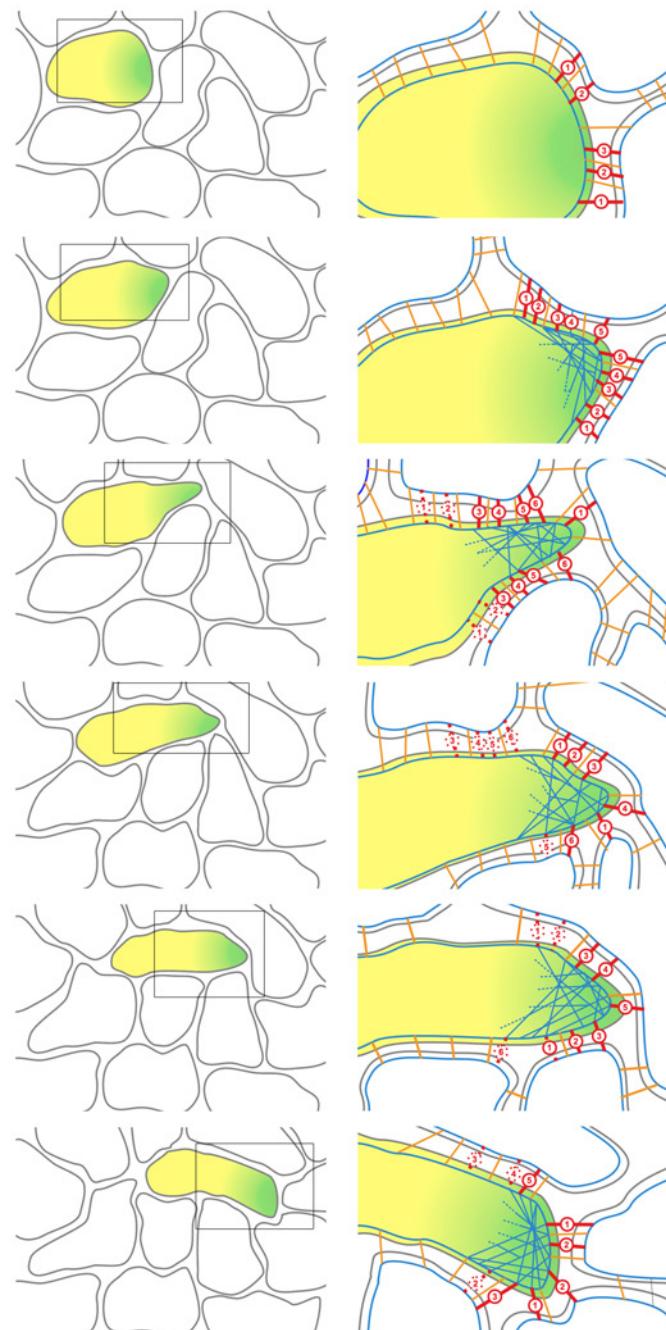


Figure 3.31: Cellular protrusion illustrated at the subcellular level. The cell cytoskeleton is in blue, the cell membrane in grey, the thin orange edges are the regular adhesion bridges and the thick red edges are focal adhesion points, related to the protrusion. The gradient of color in the cell interior reflects the polarization axis. New focal adhesion points appear at the tip of the cell, move back, then disappear as the cell moves forward.

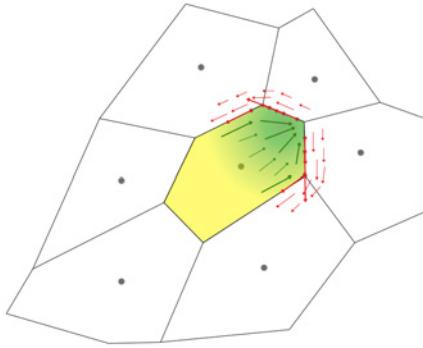


Figure 3.32: Idealized representation of the protrusive activity in MECAGEN. Green arrows represent the cell interior flow and red arrows the cell neighborhood flow, as the central cell exerts a protrusion over its surroundings.

positioned relative to i in the same general direction as \vec{U}_i (green pie-slice-shaped domain, comprising here cells j and k):

$$\mathcal{N}_i^{t+} = \left\{ j \in \mathcal{N}_i^t : \frac{\vec{X}_j - \vec{X}_i}{\|\vec{X}_j - \vec{X}_i\|} \cdot \vec{U}_i \geq \eta \right\}$$

A threshold value η controls the relative size of the protrusion (the angle of the pie slice). Similarly, we denote by \mathcal{N}_i^{t-} the sublist of neighbors which share a surface of contact on the “negative” pole of the cell, i.e. away from the polarization vector. This “negative” neighbor is used in the case of an opposite monopolar or a bipolar protrusion behavior.

In the regular monopolar case illustrated here, for each neighbor j belonging to \mathcal{N}_i^{t+} , a pair of equal and opposite forces contribute to the motion of both i and j : an intrinsic force $\vec{F}_{ij}^{A,\text{int}}$ (larger dashed green arrow) and its simultaneous and exact counterpart $\vec{F}_{ji}^{A,\text{ext}}$ (larger dashed red arrow), such that $\vec{F}_{ji}^{A,\text{ext}} = -\vec{F}_{ij}^{A,\text{int}}$. The common axis of these forces is designed to roughly emulate the profile of the surface of contact that can be seen in the polygonal representation of Fig. 3.32 (but not in the disc-particle representation of Fig. 3.33). It is a linear combination of the unitary protrusion axis \vec{U}_i and its orthogonal complement $\vec{U}_i^{\perp j}$ that passes through j , i.e. is calculated by:

$$\vec{F}_{ij}^{A,\text{int}} = f^A \cdot A_{ij} \cdot (\cos(\nu) \vec{U}_i + \sin(\nu) \vec{U}_i^{\perp j}),$$

where the angle ν tunes the profile of the surface of contact (small dashed arrows

resulting in the larger dashed arrows) and coefficient f^A tunes the intensity of the force. The angle could be precisely calculated as a function of the angular position of each neighbor, but we deemed such sophistication unnecessary and opted for a simpler scheme where it is kept constant. We chose here the larger of the two angles that a right triangle with proportions 3-4-5 makes, i.e. $\arctan(4/3) \approx 53$ degrees or 0.93 radians. Experimenting with various angle values around 45 degrees, we observed that they made no significant difference. These two parameters ν and f^A will be used especially in Chapter 5 and Case Study 8.5.

Let us consider for now that only cell i exerts a monopolar protrusive activity on its neighborhood (**Fig. 3.33B**). Each neighbor j belonging to \mathcal{N}_i^{t+} contributes a pair of forces $\vec{F}_{ij}^{A,\text{int}}, \vec{F}_{ji}^{A,\text{ext}}$ described above. Therefore, if cell i has multiple neighbors in its polar domain (here, j and k), the intrinsic forces $\vec{F}_{ij}^{A,\text{int}}$ and $\vec{F}_{ik}^{A,\text{int}}$ resulting from each neighboring interaction (both dashed green arrows) are added to produce the final behavioral force $\vec{F}_i^{A,\text{int}}$ generated by cell i 's protrusion (solid green arrow; note that it is not necessarily parallel to \vec{U}_i as we are in 3D).

Then, the same scenario could also happen around cell j , i.e. this cell too could display protrusion, thus create an equivalent set of forces around it (**Fig. 3.33C**, symmetrical with respect to Fig. 3.33B). Finally, combining both protruding activities from i and j , each cell in the neighborhood can be the site of both intrinsic and extrinsic forces (**Fig. 3.33D**): the former come from its own protruding activity (solid green arrows), while the latter from the protruding activity of its neighbors (dashed red arrows). In this particular illustration, the third cell k is not protruding, thus its own sum of active forces is only composed of extrinsic components (dashed red arrows) coming from neighbors i and j . The sum of the forces on i (resp. j and k) yields the net “active” force on this cell, which corresponds to \vec{F}_i^A (resp. \vec{F}_j^A and \vec{F}_k^A) in its motion equation at the bottom of Section 3.2.1 (but is not shown in Fig. 3.33D). Here, i has one ‘int’ and one ‘ext’ component, j also has one ‘int’ and one ‘ext’ component, while k has 0 ‘int’ and 2 ‘ext’ components. Each net active force leads a cell to move alongside its neighbors and pass through.

Finally, at the scale of the whole embryo, the net global force resulting from this complex field of local intrinsic/extrinsic active forces is zero, due to their mutual compensation: $\vec{F}^A \equiv \sum_i \sum_{j \in \mathcal{N}_i} \vec{F}_{ij}^A = \sum_i \vec{F}_i^A = \vec{0}$. This is a reasonable expectation, as the relative movements of cells with respect to each other (protrusion, construction, migration, etc.) should not have the effect of displacing the embryo but only reshaping it.

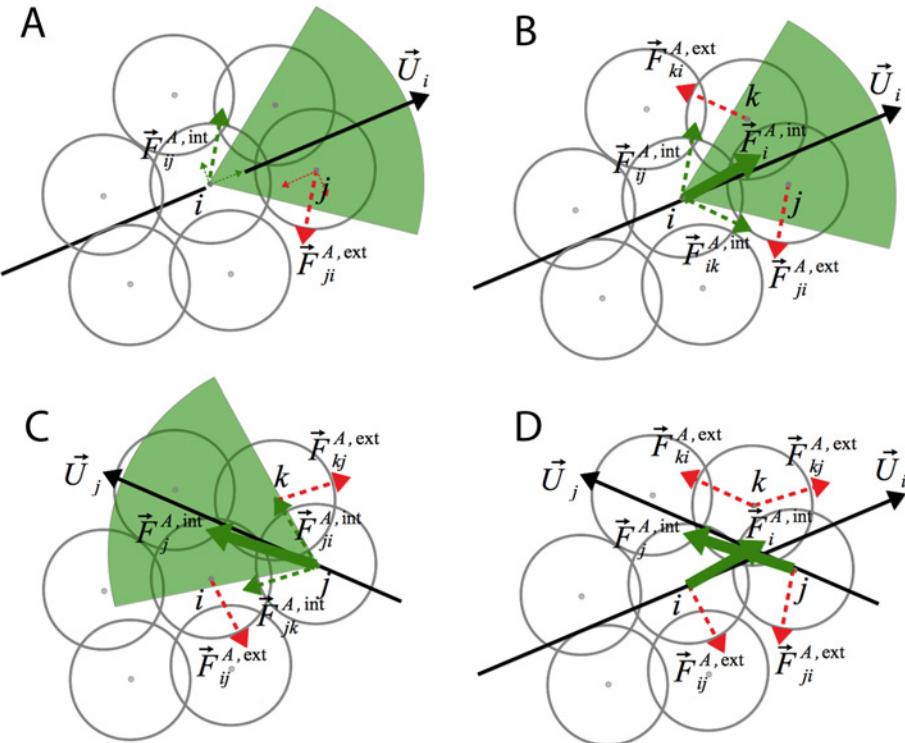


Figure 3.33: Schematic representation of active protrusive activity in MECAGEN. See text for details. This illustrates the formalization of the idealized mechanism of Fig. 3.32 in the particle-based framework. (A,B) The polar domain (green slice) of cell i , denoted by \mathcal{N}_i^{t+} , contains two neighbor cells j and k , over which i exerts a protrusive force. “Intrinsic” forces are shown in green and extrinsic forces in red. (A) Highlighting one pair of opposite forces. (B) Forces produced by i ’s activity. (C) Forces produced by j ’s activity if j , too, happens to be protruding. (D) The net resulting “active” forces (not shown) are obtained by adding the net resulting ‘int’ and ‘ext’ arrows on each cell.

CHAPTER 4

MECAGEN Model of Molecular and Genetic Regulation and Signaling

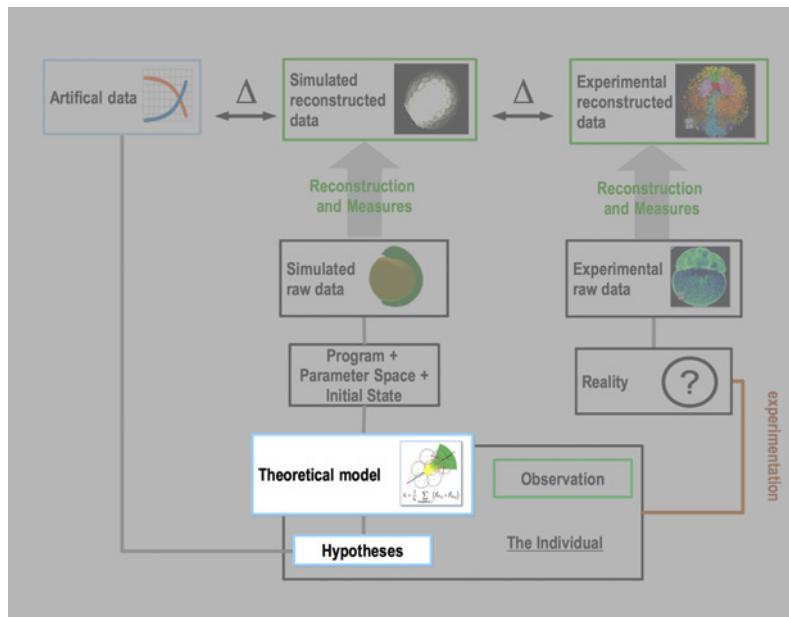


Figure 4.1: Situation of Chapter 4 in the methodological workflow.

The goal of this chapter is to briefly explain the principles of gene regulatory networks (GRNs), describe the components of GRN models, and give examples (Fig. 4.1). In the present work, our particular objective was to design a simple and easily computable model of the molecular and genetic interactions that occur during development. This domain is a subject of intense research [235][236][237], at the crossroads between bioinformatics, systems biology and chemical kinetics; nonetheless, we believe that relevant insight can already be gained by using simple rules. We articulate our model around three types of rules:

- rules driving the dynamics of intracellular gene/protein reactions

- rules driving the dynamics of cellular secretion and transduction
- rules driving the dynamics of extracellular reactions, transport and diffusion

These rules are expressed in a chemical kinetic framework by ordinary differential equations (ODEs) of the type $dp/dt = f(p, g, q, r)$, where p represents protein concentrations, g gene expression level, q external ligands, and r membrane receptors. Extracellular reactions, transport and diffusion of ligands are also taken into account via partial differential equations (PDEs) involving $\partial q/\partial t$ and fluxes $\vec{J} = -D\vec{\nabla}q$.

To facilitate the specification of the parameters of molecular and genetic regulation, we interface the model with the BioTapestry software developed by W. Longabaugh and H. Bolouri at the Eric Davidson Lab, Caltech [188][238] (Fig. 4.2). This software has been adopted by the developmental biology community as the standard tool to visualize gene regulatory networks. We use it to draw our own GRN and add different module parameters, then export it as an XML file used as an input into our model.

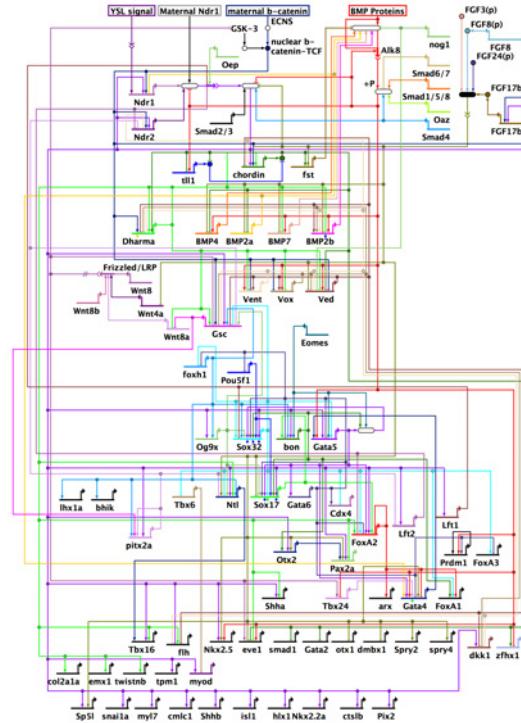


Figure 4.2: Example of BioTapestry representation of a developmental gene regulatory network (GRN) in the zebrafish. Figure from Chan et al. (2009) [44].

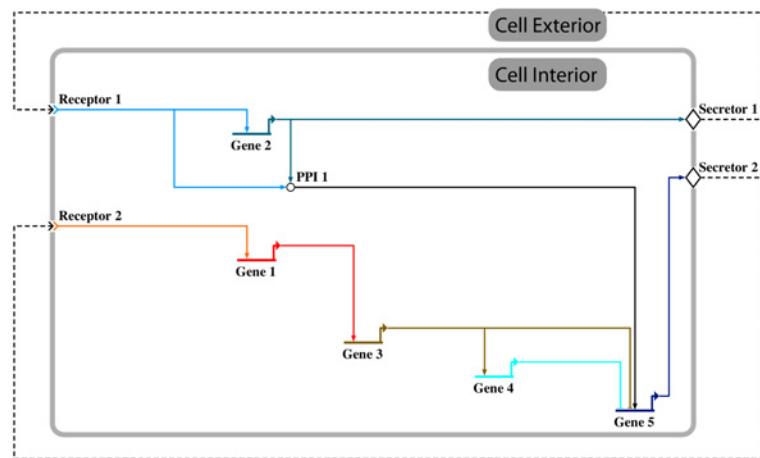


Figure 4.3: Example of abstract gene regulatory network (GRN) produced for this study using the BioTapestry software. This GRN illustrates the typical molecular/genetic regulation and signaling components used in MECAGEN: protein-protein interactions (“PPI”), genes and their cis-regulatory elements (“Gene X”), signal secretion module (“Secretor X”), signal transduction module (“Receptor X”). The solid arrows in the cell interior represent intracellular proteins; the dashed arrows outside the cell membrane represent extracellular ligands. Practically, the topology and parameters of the GRN are all specified by the user of BioTapestry, then an XML file is saved and can serve as an input into the model.

4.1 Intra-Cellular Gene/Protein Reactions

To simplify the interactions inside the cell, we restrict the variables to real-valued quantities $\mathbf{p} = \{p_a\}$ representing the different proteins $\{P_a\}$, with $a = 0 \dots N_p - 1$, and binary states $\mathbf{g} = \{g_b\}$ representing the expression levels of genes $\{G_b\}$ involved in the GRN, with $b = 0 \dots N_g - 1$. Note that p_a can stand either for an actual “number” of molecules or for a “concentration” of molecule type (i.e. the number divided by the volume of the cell). This can be justified by assuming that at every stage of the early development each cell occupies about the same volume as the other cells, and also considering that the extracellular space belongs to the Voronoi territories of the cells (i.e. that the abstract “borderline” between two cell territories is running along the middle of their interstitium). Moreover, from one stage to the next, when a cell divides, the quantities of various protein types that it contains are roughly divided by two, like its volume, thus the concentration levels can be considered continuous functions of time even through divisions and stage transitions.

The role of RNA is bypassed here, as transcription and translation are construed as a single process. Thus in the following, the term “protein” denotes a unique molecule type synthesized by a corresponding gene, whether it is actually a protein or mRNA. Then, this “protein” may either react in a protein-protein interaction, or act as a transcription factor reacting on a gene regulatory element.

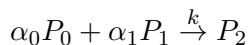
Here, each gene G_b produces a single protein type $P_{a=b}$, but on the other hand there can be more protein types than genes, i.e. $N_g \leq N_p$. This is because the GRN model only covers a small part of the whole genome and some proteins present in the cell can originate from extracellular signaling by neighboring cells. Proteins synthesized internally are in the same order as the genes in the interval $[0, N_g - 1]$, then proteins of external origin are indexed in the interval $[N_g, N_p - 1]$, if it is not empty. Protein concentrations \mathbf{p} evolve according to three rules:

- protein-protein interactions (Section 4.1.1)
- synthesis by encoding genes (Section 4.1.2)
- degradation by the molecular environment (Section 4.1.3)

Conversely, gene activities are regulated by the proteins via Boolean functions representing a logical combination of promoters and repressors (Section 4.1.4).

4.1.1 Protein-Protein Interactions

Here we consider only protein-protein reactions involving two reactants and one product, generally expressed by:



where k is the rate coefficient. The corresponding rate equation reads:

$$\frac{dp_2}{dt} = k p_0^{\alpha_0} p_1^{\alpha_1}$$

Further simplifying assumptions involve uniform stoichiometric coefficients $\alpha_0 = \alpha_1 = 1$, and one of the two protein reactants in much greater concentration than the other, e.g. $p_0 \gg p_1$, in which case the rate equation boils down to a first-order type:

$$\frac{dp_2}{dt} = -\frac{dp_1}{dt} = k' p_1$$

where $k' = kp_0$ is the new “pseudo-coefficient” of the variation. Naturally, this very simple schema has important limitations. Yet, when applied to many nodes of a large network of molecular interactions inside each cell, and also combined with the other regulatory reactions described below and spatial diffusion by cell-to-cell signaling, it can already give rise to complex spatiotemporal dynamics in cellular tissue.

4.1.2 Synthesis by Encoding Genes

We assume here that if a gene G_b is active, i.e. $g_b = 1$, the concentration p_b of the protein P_b that it encodes simply increases with a constant rate γ_b , which is characteristic of the gene:

$$\frac{dp_b}{dt} = \gamma_b g_b$$

4.1.3 Degradation by the Molecular Environment

In addition to specific protein-protein interactions and gene-to-protein synthesis, proteins P_a are also degraded by various molecular elements present in the cell. We model this process by another simple equation, based on a constant degradation coefficient κ_a characteristic of the protein:

$$\frac{dp_a}{dt} = -\kappa_a$$

Up to this point, combining the three laws above, the extended protein equations read:

$$\begin{cases} \frac{dp_1}{dt} &= -k' p_1 + \gamma_1 g_1 - \kappa_1 \\ \frac{dp_2}{dt} &= +k' p_1 + \gamma_2 g_2 - \kappa_2 \end{cases}$$

(assuming that they are both synthesized by genes, i.e. $N_g \geq 2$). In the sections below, we will add more terms coming from transduction and diffusion mechanisms, which represent the communication with the cell’s exterior.

4.1.4 Cis-Regulatory Boolean Model of Gene Expression

The activity of a gene G_b is enhanced by the “presence” of certain promoting transcription factors (TFs, subsumed under the term “protein” here) and/or the “absence” of certain repressing TFs. Both types of TFs bind to the cis-regulatory sites of the gene, which are regions of the DNA near the gene sequence. As multiple TFs P_a can potentially participate in the regulation of a cis-regulatory module G_b , something we denote by $P_a \curvearrowright G_b$, Boolean logic is a well-suited schematization of these interactions [239]. To minimize the number of variables, we represent the potential participation of a TF combined with its effective presence/absence in a regulatory module by a unique matrix of Boolean variables $\Gamma = \{\Gamma_{ab}\}$, which is globally indexed by $a = 0 \dots N_p - 1$ and $b = 0 \dots N_g - 1$, and depends on concentration levels \mathbf{p} :

$$\Gamma_{ab}(t) = \begin{cases} 1 & \text{if } P_a \curvearrowright G_b \text{ and } p_a(t) \geq \theta_{ab} \\ 0 & \text{if } P_a \curvearrowright G_b \text{ and } p_a(t) < \theta_{ab} \end{cases}$$

where the θ_{ab} parameters are concentration thresholds. The activity of gene G_b is then determined by the Boolean output of a logic function f_b , which is a combination of the Boolean operators AND, OR, and NOT:

$$g_b(t) = f_b(\mathbf{\Gamma}(\mathbf{p}(t)))$$

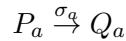
For example, if f_b is a pure AND operator, then all promoters must be present and all repressors absent for the gene to be activated. If it is a pure OR operator, then a single promoter suffices to activate G_b .

4.2 Signal Secretion and Transduction Modules

Cells in the developing embryo communicate through various means. One of the most common mechanisms are the secretion (typically by exocytosis) and the transduction (via receptors) of various molecules, such as proteins or metabolites, out of their physical domain through the cell membrane. The interfacing module connected to the GRN that internalizes and externalizes these molecules, globally denominated *ligands*, is presented in this section.

4.2.1 Signal Secretion Module

Ligands can be externalized from the cellular domain by means of secretion. A gene output that is connected to the *signal secretion* module in the GRN sends a certain quantity of its synthesized ligand P_a into the space between cells called “interstitium”, creating a concentration q_a of externalized ligand Q_a (which denotes the same molecule type as P_a , but outside the cell membrane) with a rate coefficient σ_a characteristic of ligand a . This can be represented by the reaction:



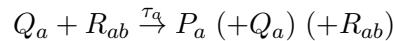
and the rate equations:

$$\frac{dq_a}{dt} = -\frac{dp_a}{dt} = \sigma_a p_a \equiv s_a$$

where s_a denotes this rate of secretion. Some externalized ligands diffuse to great distances compared to a typical cell size, while others remain attached to the cell membrane and affect only neighboring cells. Both scenarios are treated in the section below about extracellular dynamics.

4.2.2 Signal Transduction Module

Conversely, an extracellular signal is transduced into an intracellular protein through a *signal transduction* module. Three molecular actors are involved here: a receptor protein R_{ab} on the membrane, the extracellular ligand Q_a binding to the receptor, and the transduced protein downstream of the receptor P_b , which may or may not be the same molecular type as the ligand, i.e. $a = b$ or $a \neq b$ in our indexing of proteins. The transduction module is active simply if its receptor is present in the cell membrane. The receptor and ligand may or may not be consumed during this process: the receptor molecule can either stay in the membrane or disappear (e.g. internalized at the same time), while the ligand molecule can either be recycled or disappear from the interstitium (e.g. internalized or degraded). The following generic reaction summarizes these possible scenarios:



and the kinetic equations of the various molecular actors are:

$$\begin{cases} \frac{dp_a}{dt} = \tau_a \rho_{ab} q_a = \tau'_a q_a \\ \frac{dq_a}{dt} = -\varepsilon_a \frac{dp_a}{dt} = -\varepsilon_a \tau'_a q_a \equiv d_a \end{cases}$$

where ρ_{ab} is the concentration of receptor in the membrane, assumed much greater than q_a i.e. approximately constant, ε_a is a binary parameter equal to 1 if the ligand is consumed during the transduction process, and d_a denotes the rate of transduction. In sum, taking into account the effects of both secretion and transduction, the total variation of external ligand Q_a is given by:

$$\begin{aligned} \frac{dq_a}{dt} &= s_a + d_a \\ &= \sigma_a p_a - \varepsilon_a \tau'_a q_a \end{aligned}$$

4.3 Extra-Cellular Reactions, Transport and Diffusion

Various detailed models of the spatial configuration of the intersitium have been elaborated (e.g. [240]), but we prefer using here the abstract graph of neighborhood relationships (derived from a Delaunay triangulation), which was described in section 3.2.2, to serve as the infrastructure of ligand diffusion. Ligands diffuse in the interstitial regions between cells, delimited by their membranes, and the neighborhood graph connecting the centers of the cells is the “dual” representation of this space. This network also offers a spatial representation of the embryo that is robust with respect to the deformation of the multicellular assembly.

The macroscopic dynamic describing the diffusion of molecules is based on Fick’s law. It states that the ligands move from regions of high concentration to regions of low concentration with an amplitude proportional to the spatial gradient of the concentration. Generally, the flux \vec{J}_a measuring the quantity of extracellular ligand Q_a that passes through a small section of space during a small time interval is given by:

$$\vec{J}_a = -D_a \vec{\nabla} q_a$$

where D_a is the diffusion coefficient of the ligand and $q_a = q_a(x, y, z)$ its concentration field. In our network of cells, Q_a flows “on the edges” between each node i and its neighbors j (in one direction or the other). Denoting by $q_{a,i}$ the concentration of Q_a localized near the surface of cell i , and by $\vec{J}_{a,ij}$ the flux of Q_a between i and j , we can write the discrete approximation:

$$\vec{J}_{a,ij} = -D_a \frac{q_{a,j} - q_{a,i}}{r_{ij}} \vec{u}_{ij}$$

where r_{ij} is the distance between i and j , and \vec{u}_{ij} the unit vector from i to j (Fig. 4.4). Note that this expression is invariant by reversal of direction, i.e. $\vec{J}_{a,ij} = \vec{J}_{a,ji}$: it yields the same absolute gradient vector irrespective of the viewpoint, which is consistent with the existence of a unique concentration field of Q_a between i and j .

The temporal evolution of the concentration is determined by the *continuity equation*, which is a local form of conservation law. The divergence theorem gives the integral form of the continuity equation, applied on the volume of the cell. Its continuous expression reads:

$$\frac{\partial q_a}{\partial t} + \iint_O \vec{J}_a \cdot d\vec{A} = s_a + d_a$$

where $d\vec{A}$ is the normal vector of the closed surface of the cell, s_a is the “source” term corresponding to the rate of ligand produced by secretion, and d_a is the “sink” term corresponding to the rate of extracellular ligand Q_a disappearing by the transduction activity of the cell. Finally, the discrete approximation of this closed surface

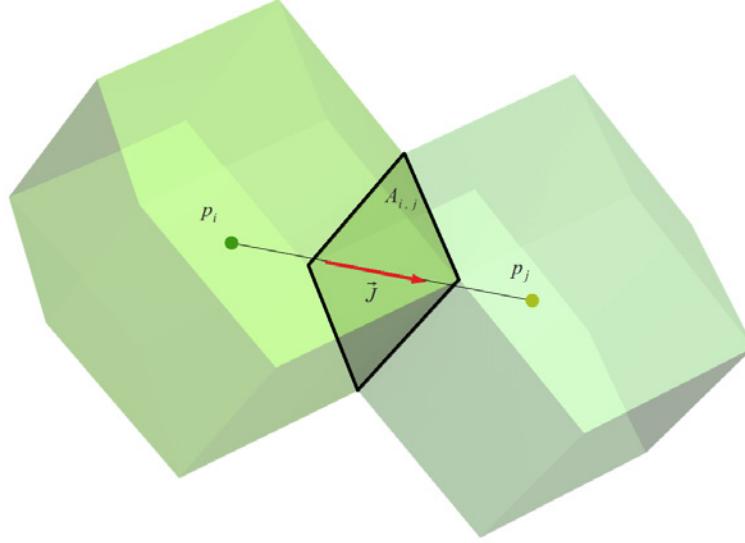


Figure 4.4: Schema of the flux of ligand. Here, \vec{J} represents $\vec{J}_{a,ij} = \vec{J}_{a,ji}$, the flux of ligand Q_a between the cellular volumes occupied by neighboring particles i and j .

integral is based on the topological neighbor list \mathcal{N}_i^t (defined in Section 3.2.2) as follows:

$$\begin{aligned}\frac{\partial q_{a,i}}{\partial t} &= - \left(\sum_{j \in \mathcal{N}_i^t} \|\vec{J}_{a,ij}\| A_{i,j} \right) + s_{a,i} + d_{a,i} \\ &= -D_a \left(\sum_{j \in \mathcal{N}_i^t} \frac{A_{ij}}{r_{ij}} (q_{a,j} - q_{a,i}) \right) + \sigma_a p_{a,i} - \varepsilon_a \tau'_a q_{a,i}\end{aligned}$$

4.4 Illustration on Artificial GRN Motifs

This section offers a glimpse of the possibilities of our molecular/genetic regulation and signaling model through simple idealized examples. We follow here E.H. Davidson's paper "Emerging properties of animal gene regulatory networks" [241], which describes various small GRN subcircuits and show their involvement in embryonic development. This section focuses on one of them: the *double negative gate*.

This first example is a part of the sea urchin embryo's GRN [242][243], in which it allows the activation of a series of genes in a specific region of the embryo under the control of localized expression (represented by protein X in Figs. 4.5-4.7). The interesting feature of this circuit is that protein X does not directly promote the

set of downregulated genes Target1 and Target2, but rather inhibits an inhibitor of these genes (Repressor1 and Repressor2). The net effect is that target genes are expressed in a particular region of the embryo and shut down everywhere else.

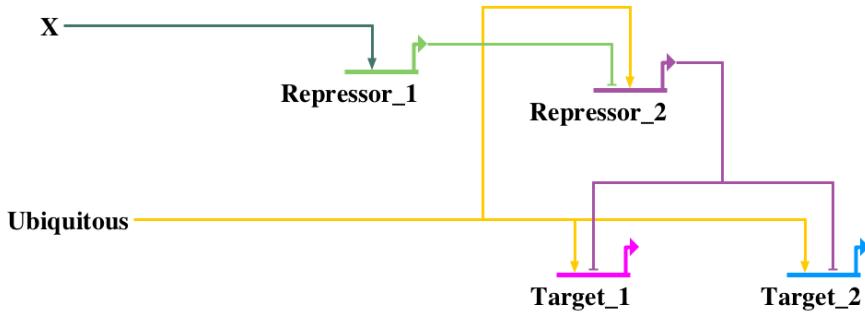


Figure 4.5: “Double negative gate” GRN subcircuit. We generated this network map with the BioTapestry software from the input file of our experimental model.

We illustrate the dynamics of this particular motif in an artificial cell population comprising 4,886 cells laid out in a two-cell-deep planar configuration. The cells are immobile as they are in an equilibrium state and no active forces are present. In the beginning, a self-descriptive “Ubiquitous” protein in all cells activates at the same time the target genes Target1 and Target2 *and* their repressor Repressor2, so that only the Repressor2 encoded protein is expressed ubiquitously (Fig. 4.6A, Fig. 4.7A). At a later point in time, protein X is introduced in one region of the cell population by artificially switching its concentration rate in these cells to a constant value of 0.1 unit per time step. In parallel, all proteins have a similar degradation rate of 0.99 unit per time step so that the X concentration tends toward an equilibrium quantity of 10 units. As soon as the X concentration exceeds the threshold $\theta_{X,Rep1}$ of the cis-regulatory element of Repressor1 ($\theta_{X,Rep1} = 1$), the Repressor1 protein starts to be produced at a rate of 0.1 unit per time step (Fig. 4.6B, Fig. 4.7B). Once the concentration of protein Repressor1 exceeds, in turn, the threshold on the cis-regulatory element of Repressor2 ($\theta_{Rep1,Rep2} = 9$), the gene state of Repressor2 switches to 0 via the AND Boolean function relating the Repressor1 and Ubiquitous TFs. Finally, the concentration of Repressor2 protein, which is no longer produced by its encoding gene, decreases by degradation. Once it passes below the concentration threshold on the sites of genes Target1 and Target2 ($\theta_{Rep2,Tar1} = 9$ and $\theta_{Rep2,Tar2} = 9$) via another AND operator, the target genes start being expressed in the region of the X protein (Fig. 4.6C, Fig. 4.7C). The time evolution of all proteins involved in region of X is shown in Fig. 4.6, and their

spatial map in Fig. 4.7 and Movie 4.1.

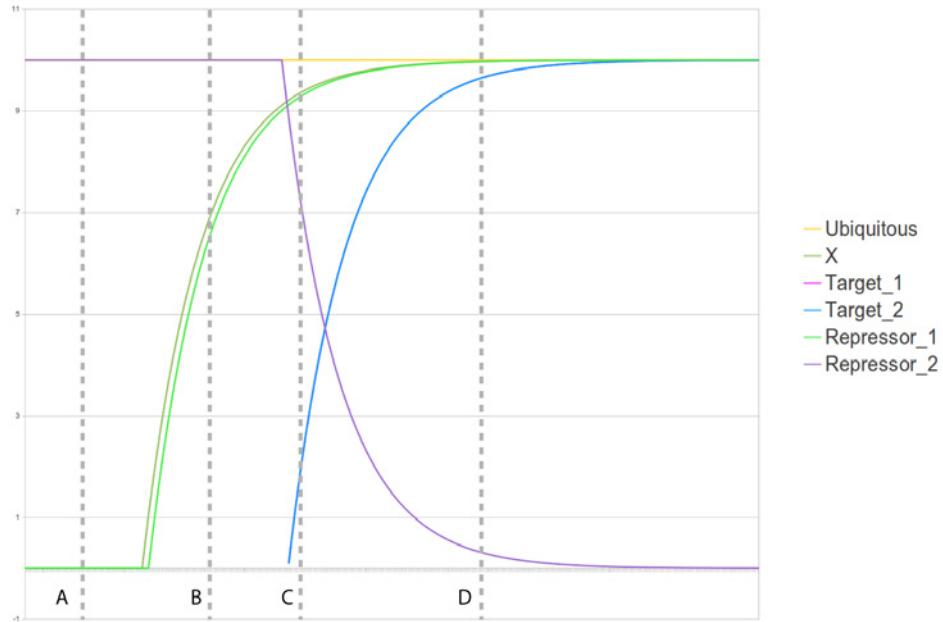


Figure 4.6: Evolution of the proteins concentration in the region where X is expressed. See text for the explanation of the curves' profile. The tagged vertical bars A, B, C, and D indicate the timing of the snapshots shown in Fig. 4.7. The curve of Target1 (pink) is hidden by the curve of Target2 (blue) as their dynamics is exactly the same.

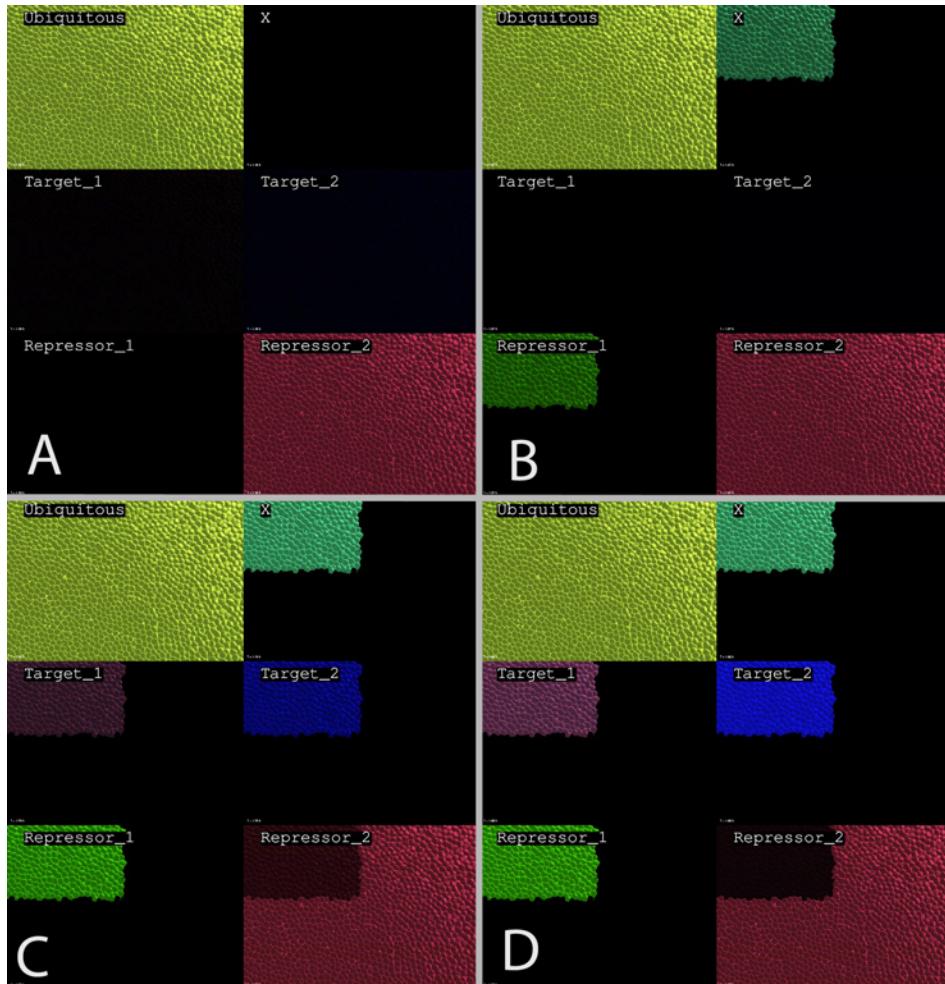


Figure 4.7: Spatial evolution of the protein concentration in the double negative gate experiment. The letters correspond to the bars of Fig. 4.6. Each image is composed of 6 simultaneous views of the simulated cell population. Colors represent protein types and shades represent their concentrations (dark for low, bright for high). The top left corner of each image is the region where protein X is artificially secreted.

CHAPTER 5

MECAGEN Model of Mechanic-Genetic Coupling

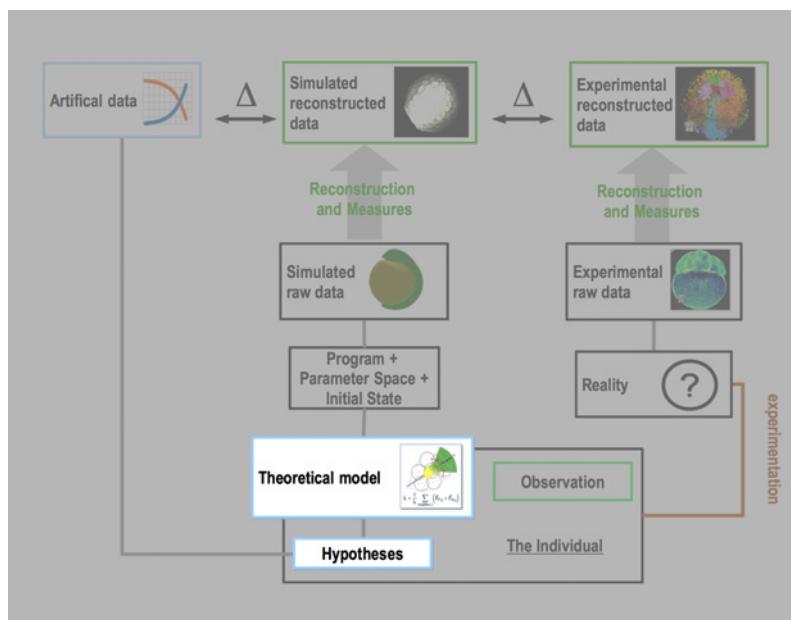


Figure 5.1: Situation of Chapter 5 in the methodological workflow.

This last of the three modeling chapters (Fig. 5.1) lays out the first steps toward building a complete morphogenetic platform integrating mechanics (from Chapter 3) and genetics (from Chapter 4), as envisioned by the MECAGEN project. It proposes a simplified “cell behavior ontology” (CBO) and relates *cell states*, determined by the concentrations of certain proteins, to *cell behaviors*, determined by biomechanical parameters. In this dissertation, we describe this coupling in two different “modes”: basically, what should be done and what has been done so far:

- Section 5.1: An “ideal” coupling mode, the proper MECAGEN model of coupling, which plugs the biomechanical module and the molecular/genetic regulation and signaling module into each other. This coupling goal has also been driving certain design choices in both parts from the beginning. However,

these principles will only be introduced here as, in its current state, the model has not been tested yet.

- Section 5.2: A “practical” and simplified model of cell behavior specification, inspired by Waddington’s epigenetic landscapes, which bypasses the actual molecular and genetic processes and replaces their kinetic equations (defined in Section 4.2) with predefined cell states, or *cell types*, and “lookup-table” parameters with one set of values per pair of cell types. Only the diffusion dynamics (of Section 4.3) will be kept. This continuous-to-discrete transformation has been commonly practiced [244] and will allow us here to test the mechanical hypotheses introduced in Chapter 3.

Finally, Section 5.3 illustrates our practical model by an artificial challenge of cell sorting and rearrangement, and a study of single-cell migration.

5.1 Toward a Cell Behavior Ontology

Biological systems modeling requires the choice of an “*ontology*”, i.e. a set of predefined objects organized by relationships. In our domain, the appropriate ontology is a categorization of the various cell behaviors that occur during a developmental process. Among a number of initiatives, we can mention James Glazier’s Cell Behavior Ontology (CBO) initiative, which proposes an in-depth classification of heterogeneous objects and processes involved in multicellular systems, in particular development. Ontologies are generally composed of a hierarchy of *categories* reflecting the vision of the ontology designers.

In MECAGEN, we propose to introduce a simpler ontology based on the dynamical rules described in Chapters 3 and 4 above. The objective is to establish a two-way mapping between the genetic and molecular states of the cell on the one hand, and its mechanical behavior on the other hand. Ideally, all parameters of the biomechanical dynamics should be defined by the molecular state of the cell, as mechanical properties emerge from both the “structural” and “regulative” molecular microstructure. It means that any spatiotemporal specification of the mechanical behavior of a cell should be ruled dynamically by the outputs of its molecular and genetic regulation. Thus, formally, the project is about monitoring a certain subset of intracellular concentration levels in each cell: $\mathbf{p}_i = \{p_{a,i}\}_{a \in M}$ (defined in Section 4.1), where $M \subset [0, N_p - 1]$, to determine the values taken by the following biomechanical parameters:

- the coefficients w_{adh} modulating the attraction part of the “passive” relaxation forces \vec{F}_{ij}^P (defined in Section 3.2.3)
- binary switches ψ deciding whether or not to apply the “active” protrusion forces \vec{F}_{ij}^A , and

- the polarization axis \vec{U}_i of each cell i giving its direction of protrusion (both defined in Section 3.2.4)

All three types of parameters are potentially functions of \mathbf{p}_i , but the first two types can be implemented at various levels of detail. The choice is typically between assigning a different value to every pair of cells i, j , thus writing these parameters $w_{ij}^{\text{adh}} = w_{\text{adh}}(\mathbf{p}_i, \mathbf{p}_j)$ and $\psi_{ij} = \psi(\mathbf{p}_i, \mathbf{p}_j)$, or only to every pair of *cell types* $\mathcal{T}, \mathcal{T}'$ covering each a certain number of cells that share the same parameters, thus writing $w_{\mathcal{T}\mathcal{T}'}^{\text{adh}} = w_{\text{adh}}(\mathbf{p}_{\mathcal{T}}, \mathbf{p}_{\mathcal{T}'})$ and $\psi_{\mathcal{T}\mathcal{T}'} = \psi(\mathbf{p}_{\mathcal{T}}, \mathbf{p}_{\mathcal{T}'})$, where $\mathbf{p}_{\mathcal{T}}$ can be for example an average over $\{\mathbf{p}_i\}_{i \in \mathcal{T}}$.

Other cell features will be decoupled from the intracellular molecular and genetic regulation in MECAGEN, and receive fixed parameter values:

- cell cycle length
- cell volume control,

while some mechanisms will not be included at all:

- cell death
- extracellular matrix.

5.1.1 Cell Behavior Ontology

In our CBO, we draw a fundamental distinction between *mesenchymal* cell behaviors and *epithelial* cell behaviors:

- Mesenchymal cells are characterized by their motility, and we identify only one “active” behavior for them: **protrusion**. By protruding, mesenchymal cells are able to rearrange their neighborhood and/or migrate through it. Protrusion can be “monopolar” or “bipolar”, depending on whether it happens on one end or two ends of the cell (the neighborhoods \mathcal{N}_i^{t+} and/or \mathcal{N}_i^{t-} defined in Section 3.2.4). This behavior rests upon the following mechanisms:

- cell adhesion modulation (represented by the coefficient w_{adh})
- cell polarization (represented by its axis \vec{U}_i)

- Epithelial cells are characterized by a reinforcement of lateral cellular junctions, and we identify two potential “active” behaviors for them: (a) **intercalation** in the lateral plane, orthogonal to the apico-basal axis, and (b) **apical constriction**, which allows a bending of the epithelial sheet (invagination). The following mechanisms are involved:

- apico-basal adhesion
- lateral adhesion
- specification of an apico-basal polarization axis
- reinforced lateral adhesion

In the remainder of this chapter, we will treat only the case of the mesenchymal cells.

5.1.2 Mechanotransduction Input of the GRN

Recent studies have shown that some genes are upregulated by mechanical forces exerted on the cell [75]. This principle should be an integral part of the MECAGEN project, as it contributes to the “MECA-to-GEN” coupling part. Practically, it would mean adding a new module to our molecular and genetic simulation platform. This module becomes activated if the sum of the forces exerted on the cell are above a given threshold and simply increases the concentration change rate of a target protein P_a by a characteristic constant ξ_a . Thus, whatever dp_a/dt ends up being at the end of Chapter 4 (after collecting various influences from genes, other proteins, secretion and/or transduction) it would now become:

$$\frac{dp_a}{dt} \leftarrow \frac{dp_a}{dt} + \xi_a$$

5.1.3 Cell Adhesion

Cell adhesion is the mechanical phenomenon that is probably most straightforward to relate to an output of molecular and genetic regulation. The intensity of the adhesion between two neighbor cells is clearly a function of the surface densities of adhesion molecules. Assuming that adhesive molecules are uniformly distributed on the cell membrane, we define at the interface between two neighbor cells i and j the surface concentration $c_{a,ij}$ (respectively $c_{a,ji}$) of protein species P_a on the membrane of i (respectively j), and express it as a function of the internal concentrations $p_{a,i}$ and $p_{a,j}$:

$$c_{a,ij}(\mathbf{p}) = \frac{A_{ij}}{4\pi R_i^2} \cdot p_{a,i}$$

$$c_{a,ji}(\mathbf{p}) = \frac{A_{ij}}{4\pi R_j^2} \cdot p_{a,j}$$

where A_{ij} is the surface of contact between the cells (introduced in Section 3.2.2), and R_i and R_j are their radii (and notation \mathbf{p} subsumes the dependency on $\mathbf{p}_i, \mathbf{p}_j$). In addition, we assume that adhesive molecules bind individually at the cell-cell interface [245] (other modes of binding are explored in [159]), thus the

equation leading the adhesion coefficient $w_{a,ij}^{\text{adh}}$ of the relaxation force \vec{F}_{ij}^P created by the adhesion molecules P_a will be:

$$w_{a,ij}^{\text{adh}}(\mathbf{p}) = k_a^{\text{adh}} c_{a,ij}(\mathbf{p}) c_{a,ji}(\mathbf{p})$$

where k_a^{adh} is an adhesion coefficient proper to molecule species P_a . As multiple adhesive molecules types are generally involved between two neighbor cells, and assuming that each molecule exerts “homotypic” adhesion i.e. binds only with its own type, then we can sum the contributions of various P_a species as follows:

$$w_{ij}^{\text{adh}}(\mathbf{p}) = \sum_{a \in M_{\text{adh}}} w_{a,ij}^{\text{adh}}(\mathbf{p})$$

where $M_{\text{adh}} \subset M$ is the family of adhesion proteins within the set of MECAGEN coupling proteins.

5.1.4 Cell Polarization

Every “active” cell behavior exploited in MECAGEN requires a *polarization axis*. In real cells, polarization correlates with an asymmetry of intracellular molecular concentrations. In our model, however, since we made the choice of one particle per cell, there can be no spatialization of intracellular material. Thus we chose to represent this asymmetry by 3D vectors \vec{U}_i passing through the centers of the cells (Section 3.2.4). More precisely, here, a cell i can be potentially polarized by multiple mechanisms as the developmental process unfolds, corresponding to multiple “candidate” polarization axes, which we denote by $\{\vec{U}_{d,i}\}_d$ with $d = 1 \dots N_d$. Each of them is associated with a binary variable ψ_d expressing whether or not $\vec{U}_{d,i}$ contributes to \vec{U}_i . Generally, for a mesenchymal cell, only one of these flags is nonzero if the cell displays an active protrusion behavior, otherwise they are all zero and no polarization axis needs to be defined (the case of epithelial cells is more complex, with two possible polarization axes, but will not be described here to avoid cluttering the notations). The link between the $\vec{U}_{d,i}$ ’s and \vec{U}_i will be further explained in the context of cell type-to-type interactions in Section 5.2.4. For now, we keep the complete list $\{\vec{U}_{d,i}\}_d$ and describe how it is managed.

We define here three possible modes of determination for each candidate axis of polarization:

- (a) a local gradient-based or “chemotactic” mode (one of the “GEN-to-MECA” coupling links)
- (b) a cell-cell contact propagation mode, and
- (c) a force induced mode (both modes of the “MECA-to-MECA” sort).

Additionally, a default mode (d) is selected if a polarized cell has no input to trigger any of these three mechanisms. The default mode randomly re-orients the

polarization axis until another polarization mode takes over. In sum, each $\vec{U}_{d,i}$ may be determined by one of these four modes—and, within mode (a), by several possible protein types P_a . Before describing these modes in more detail, we first clarify the update algorithm of each candidate polarization axis. The new candidate axis $\vec{U}'_{d,i}$ is instantaneously calculated in one simulation step, however, it is not immediately adopted and instead is combined with the current vector $\vec{U}_{d,i}$ via a “memory” coefficient $\omega \geq 0$:

- 1. initial normalization step:

$$\vec{U}'_{d,i} \leftarrow \frac{\vec{U}'_{d,i}}{\|\vec{U}'_{d,i}\|}$$

- 2. update step:

$$\vec{U}_{d,i} \leftarrow \omega \vec{U}_{d,i} + \vec{U}'_{d,i}$$

- 3. final normalization step:

$$\vec{U}_{d,i} \leftarrow \frac{\vec{U}_{d,i}}{\|\vec{U}_{d,i}\|}$$

We now explain the four modes of polarization:

- (a) **Local gradient-based mode:** This mode expresses a classical vision of polarization, which is for example often used in *chemotactic* cell behavior, and supposes that the cell is able to detect an asymmetry of the extracellular ligand concentration in its local vicinity. In MECAGEN, this asymmetry is determined using the abstract graph of neighborhood relationships \mathcal{N}_i^t (Section 3.2.2) and the extracellular ligand quantities q_a (Section 4.2.1). This mode is associated with one particular ligand type Q_a and defines the candidate axis of polarization to be a weighted average of the neighborhood edges \vec{u}_{ij} , in which the weights are polynomial functions of the differences of ligand concentrations:

$$\vec{U}'_{d,i} = \sum_{j \in \mathcal{N}_i^t} (q_{a,j} - q_{a,i})^m \vec{u}_{ij}$$

where $q_{a,i}$ (resp. $q_{a,j}$) is the local quantity of extracellular ligand Q_a surrounding i (resp. j), and m is an integer controlling the sensibility of detection of local concentration differences (m must be odd to conserve vector directions, typically $m = 3$).

- (b) **Cell-cell contact propagation mode:** This mode represents another way to establish polarization axes in a cellular field. It is based on the notion

that the asymmetric spatialization of intracellular material in a polarized cell translates into a surfacic asymmetry, which neighbor cells can sense. Only polarized neighbors j , i.e. for which $\psi_{d,j} = 1$ influence their vicinity. Thus, the cell-cell contact propagation mode simply defines the polarization axis of cell i as the average of its polarized neighbors' axes:

$$\vec{U}'_{d,i} = \sum_{j \in \mathcal{N}_i^t} \psi_{d,j} \vec{U}_{d,j}$$

(Note that this is an average of one candidate type of axis over different cells j , in which several of the $\psi_{d,j}$ flags may be equal to 1; not to be confused with a sum over different candidate types d of the same cell, in which generally only one of these flags is nonzero.)

- (c) **Force-induced mode:** Recent studies have unveiled a novel mechanism for the propagation of polarization axis based on mechanical interactions [73][246]. Weber et al. [73] demonstrate that a mechanical traction exerted at one end of a *Xenopus* cell triggers the formation of protrusion in the opposite direction through a reorganization of internal intermediate keratin filaments. In MECAGEN, we idealize this mechanism by orienting the candidate polarization axis of a cell i in the direction opposite to the average force exerted by the neighbor cells j that are protruding on and around it:

$$\vec{U}'_{d,i} = -\frac{1}{N_{A,\text{ext}}} \sum_{j \in \mathcal{N}_i^t} \vec{F}_{ij}^{A,\text{ext}}$$

where $\vec{F}_{ij}^{A,\text{ext}}$ is the “extrinsic active” force exerted by j over i (the dashed red arrows of Fig. 3.33, Section 3.2.4), and $N_{A,\text{ext}}$ is the number of neighbors j for which this contribution is not zero (although this normalization is unnecessary as it will be taken care of by step 1. in the above update algorithm).

- (d) **Default mode:** When a protruding cell does not receive spatial cues about how to orient itself, it enters a “blebbing” state. In the MECAGEN model, such cells receive a completely random polarization axis at regular time intervals, typically 5 simulated minutes.

5.2 Waddingtonian Timeline Specification

Section 5.1 described the ideal MECAGEN coupling framework, in which the physico-chemical causal and consequential factors (inputs and outputs) of the biomechanics of embryo development are the variables of a gene/molecular regulation and signaling network, as it can be typically represented in BioTapestry. In this section, our practical objective will be to study development based on cell forces and cell types alone, bypassing the molecular and gene dynamics. As Odell et al. put it

in [219], cell forces, similarly to “clocks, morphogens, or potentials”, are the “cause” which “explains” the “morphogenetic motion”. To this aim, we decided to shortcircuit molecular and genetic interactions and, instead of relying on dynamical GRN outputs, “read out” cell behaviors directly from a novel parameter specification tool, which we call the *Waddingtonian Timeline Specification* (WTS; see Section 1.1.4 for a reminder of Waddington’s ideas about “epigenetic landscapes” and cell fates). We explain in the following subsections how the mechanical parameters of our CBO (Section 5.1) can be tuned by this WTS operator. A specific graphical user interface has also been developed for this purpose and will serve as a support for the presentation.

5.2.1 Cell Types

The first step of a WTS is to “carve the hillside” by specifying a temporal series of *cell types* $\{\mathcal{T}(i, t)\}$ that cells may adopt during the developmental process, and the transition rules among these types:

- First, we segment the timeline into *stages* through particular points in time $\{t_1, t_2, \dots\}$ at which new cell types may be introduced. For example, Stage 2 corresponds to the time interval $[t_2, t_3]$. During one stage, cells may transition from one existing type to another under the rules specified in a table (see next). The overall WTS structure can be represented as a pseudo-tree of cell types (with lateral transfers among the branches) expanding over time (Fig. 5.2).
- Then, we specify the conditions inducing a cell to change its type. We employ the classical concept of *differentiation* to name this process. In our simplified model here, differentiation depends on a Wolpertian “French flag”-like *positional information* mechanism (see review in Section 2.1.3). During stage $S = [t_S, t_{S+1})$, if the local ligand concentration $q_{a,i}$ of molecular species Q_a on cell i exceeds (or sinks below) a given threshold θ_a , then cell i changes types. Generally, the cell’s new type is a function of its current type and a *differentiation operator* D represented by a ruleset or “lookup table” predefined for each stage (Fig. 5.3), which we denote by D_S . This table is organized in modules parametrized by triplets composed of ligand species, their respective thresholds and their signs denoted ε_a (expressing whether the rule is about an exceeding or a sinking concentration, i.e. $\varepsilon_a q_{a,i} \geq \varepsilon_a \theta_a$). In summary, at any time $t \in S$, a type transition can be formally written:

$$\mathcal{T}(i, t) \leftarrow D_S[\mathcal{T}(i, t), \{(Q_a, \theta_a, \varepsilon_a)\}_{a \in M}].$$

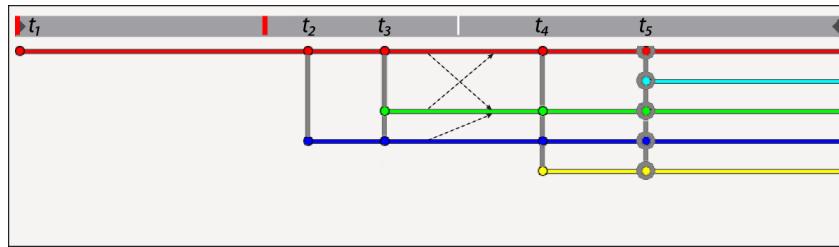


Figure 5.2: Waddingtonian timeline of cell types $\{\mathcal{T}(i, t)\}$. The grey bar on top is the time axis (oriented left to right), segmented into stages $S = [t_S, t_{S+1})$. (The thin white time tick is the current time step of the simulation and the thick red time ticks mark backup & restore points of the entire embryo state.) The colored horizontal branches symbolize the types that cells can potentially take, and correspond to the grooves of a Waddingtonian landscape (Fig. 1.3). Here, we can see 5 stages in the timeline corresponding respectively to the onset of 1 type (red), 2 types (red and blue), ..., and 5 types (red, cyan, green, blue, yellow). Differentiations (type transitions) may happen during each stage, for example shown here in Stage 3: the three thin dashed arrows correspond to the differentiation table of Fig. 5.3 below, indicating that type red may become green, green may become red, and blue may become green. In this WTS graphical interface, stages can also be selected manually to specify their parameters (as will be shown in the next figures): here, Stage 5 is selected by clicking on t_5 , which is represented by gray circles on all the type nodes.

A

3	1		
2000	1		

3	-1		
2000	0		

2	1		
2000	1		

B

Ligand ID	Sign		
Threshold	Target Cell Type		

Origin Cell Type

Figure 5.3: Visualization of the differentiation (type-transition) operator D . (A) This particular table represents the possible transitions between cell types during Stage 3 above, thus can be denoted D_3 . It contains three differentiation modules displayed in one column, one module by origin type. In each module, the current cell type $\mathcal{T}(i, t)$ is represented by the frame color, the target cell type by the color (or id) of the bottom-right block, and the ligand-threshold-sign triplet $(Q_a, \theta_a, \varepsilon_a)$ by the top-left block's color (or id), bottom-left block's value, and top-left block's value. Each cell changes its type depending on whether $\varepsilon_a q_{a,i} \geq \varepsilon_a \theta_a$. (B) The generic template of one differentiation module inside one origin type.

5.2.2 Ligand Sinks and Sources

Once the backbone of the WTS has been designed, the parameters of *ligand diffusion* must be specified. At each stage S , we define sources and sinks for each ligand through another table denoted L_S (Fig. 5.4). Each cell type can potentially secrete or absorb any ligand type Q_a .

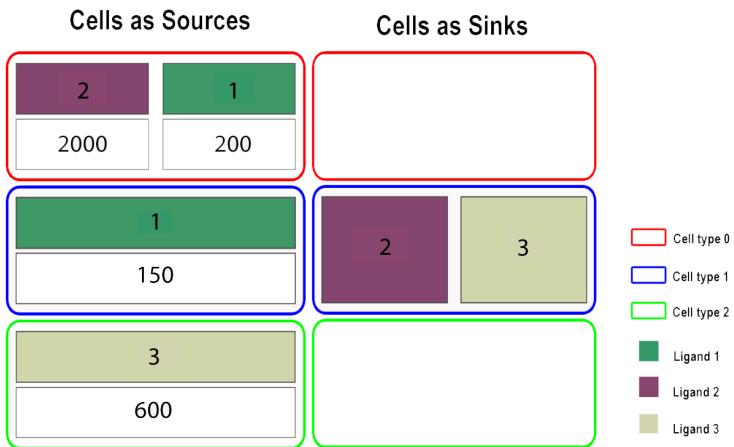


Figure 5.4: Visualization of the ligand diffusion parameters L . This is the particular table at Stage 3, thus can be denoted L_3 . Left column: Cells as sources. Each module corresponds to a current cell type (denoted by its frame color) and is composed of one or several pairs of boxes displayed vertically. The top box indicates the color (or id) of the secreted ligand Q_a . The bottom box indicates the rate s_a of ligand secreted by each cell belonging to the cell type (defined in Section 4.2.1). Thus, red cell types 0 secrete both ligand types 2 (dark green) and 1 (purple), with rates respectively 2000 and 200. Right column: Cells as sinks. Each module only specifies the type (color or id) of the ligand that the cells can absorb, if any (red and green cell types absorb none). No particular quantity is specified as we assume here that all the ligand is absorbed by the cells acting as sinks.

Additionally, the *spatial configuration* of the ligand sources must also be specified. The module used for this type of specification (not displayed here, but see two examples in Figs. 5.10 and 5.14) is adapted to the spherical geometry of the zebrafish embryo. We define, per cell type, the id of the ligand and the geometrical border of the area of secretion. Assuming a spherical embryo, an orthonormal coordinate frame is set up along the animal-vegetal (AV) axis, antero-posterior (AP) axis, and bilateral symmetry left-right (LR) axis. Then, on each axis, we define two cutoff values to extract a slice and take the intersection of all three slices to define the source region of ligand release. Several such source regions can be defined per

cell type. Thus the spatial ligand source table is essentially composed of sextuplets of cutoff coordinate values for each ligand inside each cell type module. External sources of ligands, such as the yolk, may also be added.

5.2.3 Passive Force's Adhesion Modulation

Here too, the adhesion coefficient w_{adh} of relaxation forces \vec{F}_{ij}^P is simply specified by pair of cell types, thus can be written $w_{\mathcal{T}\mathcal{T}'}^{\text{adh}}$, for $i \in \mathcal{T}$ and $j \in \mathcal{T}'$. Since passive adhesion forces are symmetrical, then $w_{\mathcal{T}\mathcal{T}'}^{\text{adh}} = w_{\mathcal{T}'\mathcal{T}}^{\text{adh}}$ and these coefficients can be represented by a triangular $\mathcal{T} \times \mathcal{T}'$ matrix P (Fig. 5.5).

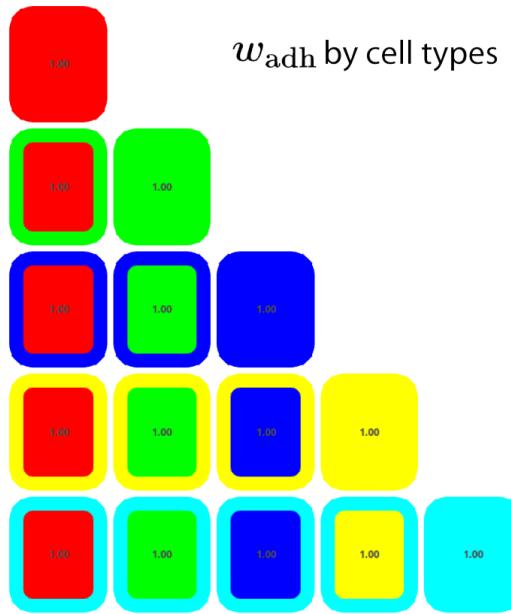


Figure 5.5: Visualization of the passive adhesion coefficients P . This is the particular table at Stage 5, thus can be denoted P_5 . The matrix is symmetrical so only half of the coefficients are displayed. The outer frame color represents cell types by row, and the inner color represents cell types by column. Here, all values of $w_{\mathcal{T}\mathcal{T}'}^{\text{adh}}$ are uniformly set to 1.00.

5.2.4 Active Cell Behavior Specification

Finally, once the timeline frame, the differentiation table D , the ligand sources/sinks table L , and the “passive” adhesion force matrix P are all defined, the “active” cell behavior can also be set up by adding other behavioral modules for a given cell type at a given stage. To this aim, an “active” protrusion table A (Fig. 5.6) is composed of modules associated to an origin cell type and containing four parameters:

- the target cell type that the protrusion is affecting

- the chosen index d among all the various candidate axes of protrusion $\{\vec{U}_{d,i}\}_d$, which can be calculated here only by one of the ligand-based mode (a) or random mode (d) (see Section 5.1.4)
- the intensity f^A of the protrusive force (length of the larger dashed arrows in Fig. 3.33, Section 3.2.4)
- a ternary value equal to +1 if the cell is monopolar in the direction of \mathcal{N}_i^{t+} , -1 if it is monopolar in the opposite direction \mathcal{N}_i^{t-} , or 0 if it is bipolar i.e. protruding in both directions.

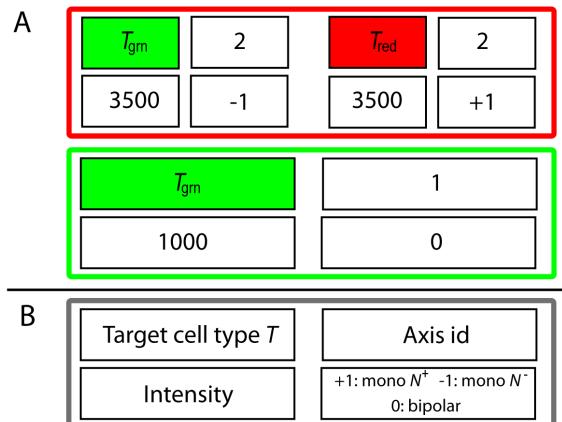


Figure 5.6: Visualization of the active protrusion coefficients A. This is the particular table at Stage 2, thus can be denoted A_2 . (A) Two cell types are involved here: red and green. In each cell type's module, protrusion is represented by four parameters. Top-left: target cell type; top-right: the chosen polarization axis id d ; bottom-left: the intensity of the protruding force f^A ; bottom-right: the type of protrusion, either monopolar+, monopolar- or bipolar. (B) The generic template of one active protrusion module inside one origin type.

In summary, The Waddingtonian timeline is a novel, yet limited, method of specifying cell behaviors through space and time. It allows a partial exploitation of the principles involved in MECAGEN but it is sufficient to start exploring the mechanical space and coupling principles. In particular, it will be used in some of the specific case studies of the Zebrafish mechanical development developed in the second part of the manuscript, especially in Section 8.5 dedicated to a study of intercalation patterns during epiboly.

5.3 Illustration on Artificial Cell Sorting

One of the most interesting aspects of the mechanical MECAGEN project is the attempt to unify the causal mechanisms of most individual and collective events involving mesenchymal cells during animal development. The major point in our study here is that *cell motility is acquired through cell protrusive activity*. In this section, we illustrate this principle through two abstract experiments, such as cell sorting (Section 5.3.1) and individual cell migration (Section 5.3.2).

Historically, spontaneous cell sorting and rearrangement is one of the multicellular phenomena that attracted the greatest number of theoretical models. Steinberg was the first to propose a fundamental biomechanical principle to explain it, which he called the Differential Adhesion Hypothesis (DAH) [81] (see brief review in Section 2.1.5). Graner and Glazier also proposed a theoretical model based on same idea expressed in the Cellular Potts framework, where the space is a grid and cell particles (pixels) change domains according the minization of a global energy term, mimicking movement [153] (see Section 3.1.1). In short, the main idea is that collective ballet of spatial cell reorganization emerges from local rearrangements at the level of cell-cell contacts via “elective affinities”.

We performed here similar sorting experiments in an abstract setting. Instead of the “temperature” parameter, an often used analogy from thermodynamics (in this particular case corresponding to an “intrinsic cell motility” coefficient [159] that controls membrane fluctuations and the sorting efficiency), we explore our own parameter ontology, centered on the orientation of the polarization axis. Our artificial model performs similarly to Cellular Potts with a different dynamics and parameter space. In the first exercise below, we illustrate this mechanism with three sorting experiments realized on a bilayer comprising about five thousand cells. The sorting mode will depend on the mechanical parameters discussed above (Section 5.1), mainly the protrusion axis and the adhesion coefficient of the relaxation force. In the second exercise, we present a simple individual cell migration experiment based on the same assumptions.

5.3.1 Cell Sorting: Revisiting a Classical Problem

In the three sorting experiments presented here, we use the same setup as in Section 4.4 about illustrating gene regulation and molecular signaling—with the essential difference that, here, cells can move but do not change types. In the general WTS framework presented above (Section 5.2.1), this means that we are looking inside one stage S with only two horizontal cell type lines, one red and one green, and no transition arrows between them. Thus there is no differentiation table D . For the spatial layout, we have again 4,886 cells in a thin 3D space delimited by two planes at a distance equivalent to four cell diameters. All figures below represent the cellular assembly from an external viewpoint that make them appear like a 2D

tissue. Initially, cells are randomly assigned one of two cell types: a “red” cell type \mathcal{T}_{red} and a “green” cell type \mathcal{T}_{grn} , creating two populations of fixed size. As usual, the forces that apply between two cells are composed of the “passive” relaxation forces \vec{F}_{ij}^P and the “active” protrusion forces \vec{F}_{ij}^A from Sections 3.2.3 and 3.2.4. Accordingly, the parameters that we will tune here are:

- the adhesion coefficient w_{adh} of the passive forces, and
- the axis of polarization \vec{U}_i of each cell (which includes the polarization mode, the force intensity f^A , and the polarity)

Cell neighborhoods will be given by the usual topological sets \mathcal{N}_i^t for passive adhesion forces (defined in Section 3.2.2) and \mathcal{N}_i^{t+} and \mathcal{N}_i^{t-} for the active protrusion forces (defined in Section 3.2.4).

Experiment I: Strong homotypic adhesion and random protrusion axis

Experiment I setup

In the first experiment, there is no ligand diffusion, thus table L of Section 5.2.2 is empty. We now adapt our view of the DAH to our framework to configure the passive adhesion table P . Both cell populations have a *strong homotypic adhesion*, i.e. a large adhesion coefficient w_{adh} for the passive forces exerted between cells of the same type, and a *weak heterotypic adhesion*, i.e. a small adhesion coefficient between cells of different type. Using our previous notations from Section 5.2.3, this means a simple two-type table P with one high value $w_{\mathcal{T}\mathcal{T}}^{\text{adh}} = w_{\mathcal{T}'\mathcal{T}'}^{\text{adh}} = 5.0$ and one low value $w_{\mathcal{T}\mathcal{T}'}^{\text{adh}} = w_{\mathcal{T}'\mathcal{T}}^{\text{adh}} = 1.0$, where $\mathcal{T}, \mathcal{T}'$ stand for $\mathcal{T}_{\text{red}}, \mathcal{T}_{\text{grn}}$ (Fig. 5.7).



Figure 5.7: Experiment I passive adhesion table P . See definition in Section 5.2.3 and text for details. In this case, homotypic adhesion is strong (5.0), heterotypic adhesion is weak (1.0).

Now, relaxation would not be sufficient in itself to trigger a sorting behavior if no active protrusive force was involved. For this purpose, we let the cells perform protrusive activity on every neighbor cell, whether they belong or not to the same population. We follow here the default mode of polarization (d) presented in Section 5.1.4., i.e. the polarization axis $\vec{U}_{d,i}$ determining which neighbor cells are inside the

polar domain \mathcal{N}_i^{t+} of a protruding cell i will receive a random orientation. In these experiments, it corresponds to an axis index $d = 3$ from a list of three candidate axes (Fig. 5.7). The other two axis indices will be explained in the experiments below.

Figure 5.8: Experiment I active protrusion table A. See definition in Section 5.2.4 and text for details.

Experiment I results

In Fig. 5.9 and Movie 5.1, we can visually observe the effective consolidation and growth of two homotypic clusters of red and green types. The sorting rate seems to tend to a plateau at some point in the simulation, however (quantity not measured). It is not clear whether the reassortment would eventually be completed, i.e. whether the two domains would reach a minimal-energy, straight boundary line between them, or how long it would take. This would require a more thorough exploration of the parameter space defined by tables P and A above. For now, the goal of this experiment was only to show a proof-of-concept.

Experiment II: Ligand-based heterotypic protrusion, planar diffusion sources

Experiment II setup

In the second experiment, the homotypic adhesion is once again stronger than the heterotypic adhesion, thus the passive adhesion table \mathcal{P} is similar to Experiment I except with different proportions (Fig. 5.11). Polarization axes, however, are not determined randomly as before; instead, it is the ligand-based chemotactic mode (a) of axis specification that is at work here (see Section 5.1.4). Two different ligand molecular species are secreted, Q_1 and Q_2 , and both types of cells \mathcal{T}_{red} and \mathcal{T}_{grn} are potential sources for these ligands with the same secretion rates s_1 and s_2 in

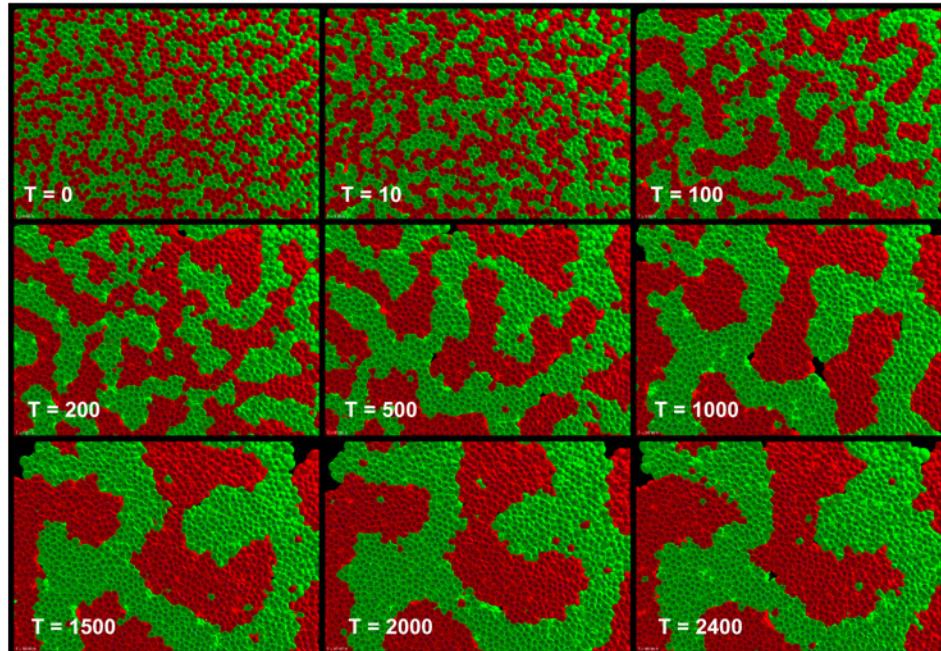


Figure 5.9: Experiment I: Strong homotypic adhesion and random protrusion axis. See text for comments. The homotypic adhesion coefficient ($w_{adh} = 5$) is higher than the heterotypic coefficient ($w_{adh} = 1$). All cells have bipolar protrusive activities with random polarization axes (mode (d) in Section 5.1). The rate of cluster size growth is decreasing over time. The time is provided through the number of simulation time steps.

the ligand table \mathcal{L} (Fig. 5.10). No cell absorbs any ligand, so there is no boundary condition on the low concentration end. On the high concentration end, however, there is a spatial constraint on where the ligand sources are located (explanation at the bottom of Section 5.2.2): two rectangular domains are predefined on the left and right border of the frame, and whenever a cell of any type enters the right (resp. left) domain, it starts secreting Q_1 (resp. Q_2).

Concerning active protrusion, the rule is that T_{red} cells only respond to the gradient created by ligand Q_1 , and T_{grn} cells only respond to Q_2 . The other parameters of the active protrusion table \mathcal{A} (Fig. 5.12) concern the protrusion force intensity f^A and the polarity, which is +1 here for both cell types, meaning that all cells protrude in the uphill direction of their gradient of choice. The net effect is that red cells orient their polarization axis toward the left border (higher q_1 concentration values) and green cells toward the right border (higher q_2). Note that, contrary to Experiment I, the protrusive behavior here is exclusively heterotypic, i.e. cells exert an active protruding force only on the neighbors that are of the other cell type (red cells protrude on the green cells only, and vice-versa). In other terms all the active mechanical interactions (described in Fig. 3.33) occur at the interface between the two populations, not within populations.

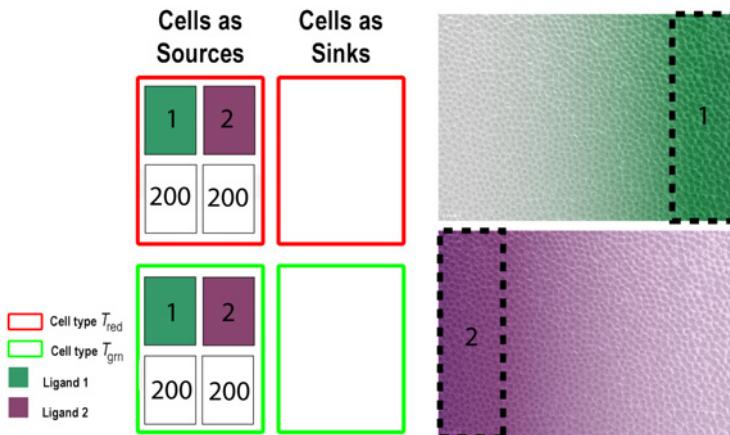


Figure 5.10: Experiment II ligand diffusion table \mathcal{L} . See definition in Section 5.2.2 and text for details. In summary, both cell types secrete both ligand types Q_1 and Q_2 with the same secretion rates $s_1 = s_2 = 200$ simulation units. They do this, however, only if they enter one of the source regions on either vertical border of the domain.

Experiment II results

Although here too cells exhibit a clear collective reassortment behavior (Fig. 5.13 and Movie 5.2), we observe again that the boundary line between the two



Figure 5.11: Experiment II passive adhesion table P . See definition in Section 5.2.3 and text for details.

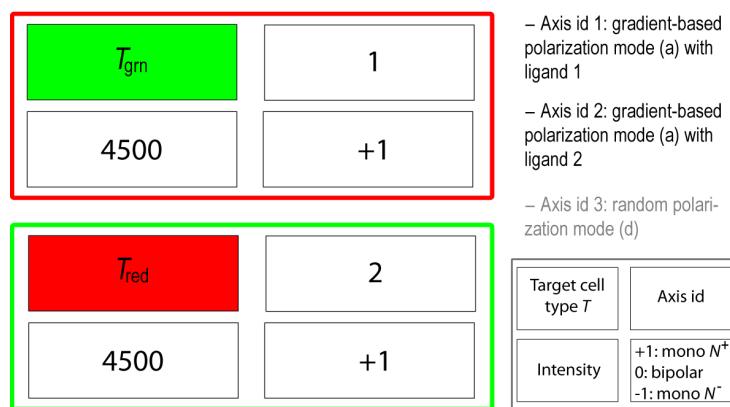


Figure 5.12: Experiment II active protrusion table A . See definition in Section 5.2.4 and text for details. In summary, protrusion behavior here is heterotypic, i.e. cells act only upon neighboring cells of the other type.

populations does not tend to become flat, as would be expected from a classical cell sorting study. In this experiment, all cell polarization axes \vec{U}_i are roughly colinear and aligned with the horizontal direction. This is because at the later stages (after time step 3500), the profile of the boundary line between the red and green populations is directly related to η , the scalar product limit determining the “positive” polar neighborhood \mathcal{N}_i^{t+} centered around \vec{U}_i (pie-slice shape in Fig. 3.33). This is due to the fact that for a green cell near the boundary line, which protrudes toward the right side of the tissue, no red cells are present any longer in its polar neighborhood. Thus the tissue dynamics reaches an equilibrium state, in which the boundary line exhibits a jagged shape. Thus the way that protrusion behavior is modeled here is not sufficient to reach the ideal straight boundary of a cell sorting phenomenon, and additional mechanisms are needed.

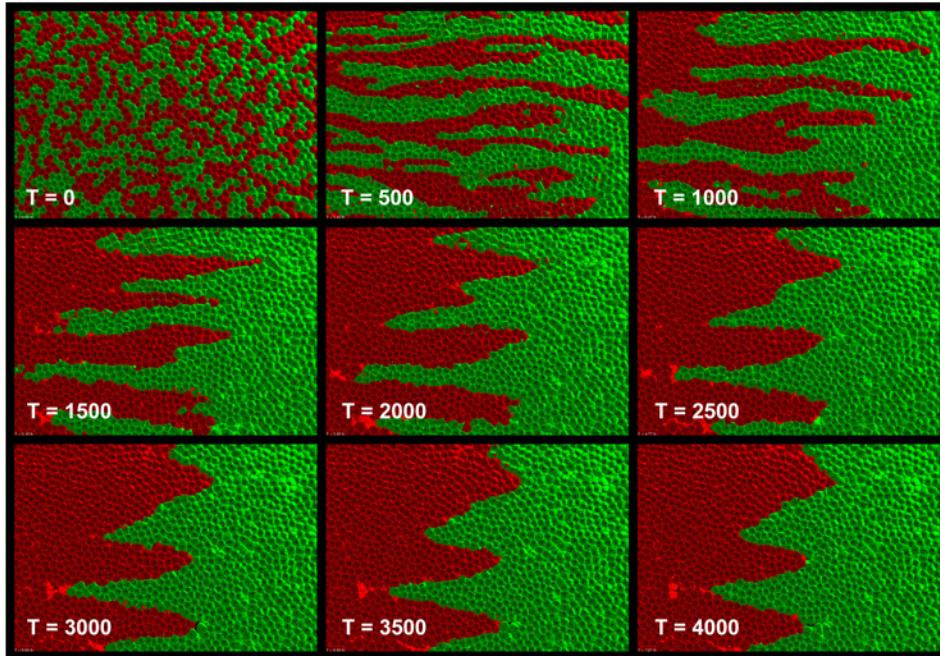


Figure 5.13: Experiment II: Ligand-based heterotypic protrusion, planar diffusion sources. See text for comments. Two hidden ligands are diffusing from the left and right borders of the cell bilayer. Green cells’ polarization axes are oriented toward the right source and red cells’ axes toward the left source. Each cell type is exerting monopolar protrusion over the other cell type (heterotypic contacts).

Experiment III: Ligand-based heterotypic protrusion, radial source

Experiment III setup

The third experiment is similar to the second one, with strong homotypic and weak heterotypic adhesion (same passive force table \mathcal{P} as in Fig. 5.11), but here a single ligand is secreted from the center of the swarm (Fig. 5.14). Moreover, protrusion polarities are in opposition here: while red cells orient their polarization axes toward the source of the ligand (uphill along the local gradient), green cells by contrast orient theirs *away* from that source (Fig. 5.15).

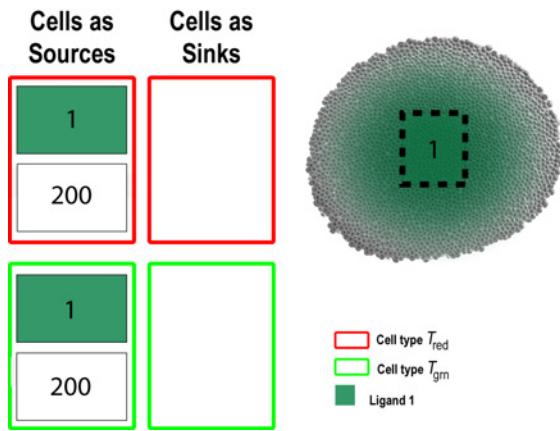


Figure 5.14: Experiment III ligand diffusion table L . See definition in Section 5.2.2 and text for details. The differences with Experiment II's table L is a unique source of ligand Q_1 and a central region of secretion creating a (quasi) radial gradient.

Experiment III results

Similarly to the previous experiment, cells reassort properly but, also for the same reasons, the boundary line between the red and green cell populations is not smooth (Fig. 5.16).

5.3.2 Individual Cell Migration

Experiment IV setup

This last experiment is meant to illustrate the generality of the protrusive behavior in MECAGEN, which can also serve as a model of individual *cell migration*. The mechanisms and parameters involved in this study are similar to the above sorting experiments. They also include passive adhesion and ligand-based chemotactic protrusion. We designed a simple scenario to qualitatively mimic the migrating

T_{gm}	1	- Axis id 1: gradient-based polarization mode (a) with ligand 1				
4500	+1	- Axis id 2: gradient-based polarization mode (a) with ligand 2				
T_{red}	1	- Axis id 3: random polarization mode (d)				
4500	-1					
		<table border="1"> <tr> <td>Target cell type T</td> <td>Axis id</td> </tr> <tr> <td>Intensity</td> <td>+1: mono N^+ 0: bipolar -1: mono N^-</td> </tr> </table>	Target cell type T	Axis id	Intensity	+1: mono N^+ 0: bipolar -1: mono N^-
Target cell type T	Axis id					
Intensity	+1: mono N^+ 0: bipolar -1: mono N^-					

Figure 5.15: Experiment III active protrusion table A. See definition in Section 5.2.4 and text for details. The differences with Experiment II's table A is that the polarization mode relies on only one type of ligand Q_1 and the green cells' monopolar protrusion is oriented in the “negative” direction.

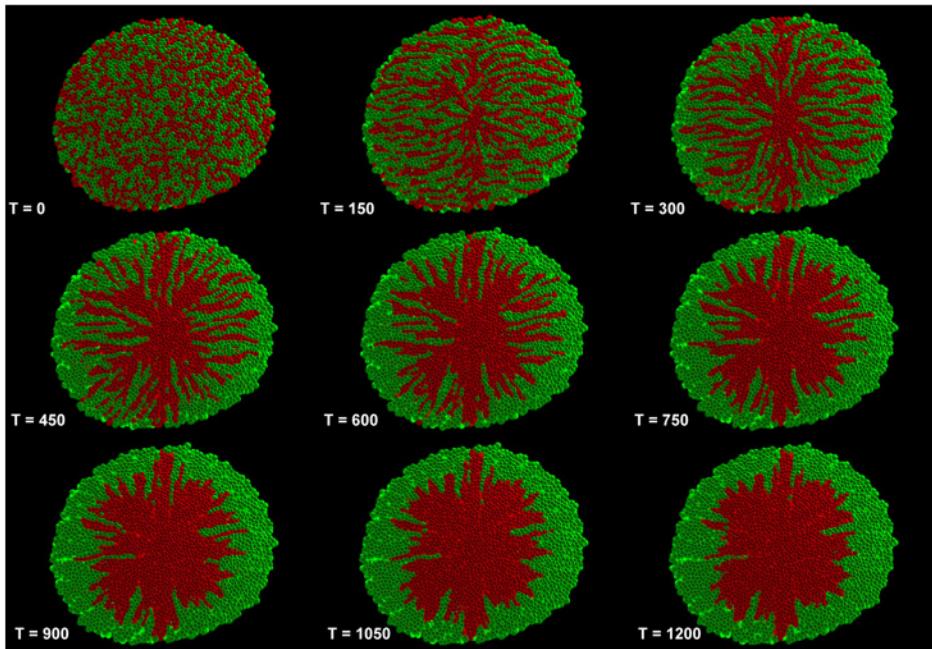


Figure 5.16: Experiment III: Ligand-based heterotypic protrusion, radial source. See text for comments. A single ligand is diffusing from the center of the cell bilayer. All cells' polarization axes are roughly radial, the red ones oriented toward the center (uphill along the gradient) and the green ones away from the center (downhill). As in Experiment II, each cell type is exerting monopolar protrusion over the other cell type (heterotypic contacts).

behavior of the germ cells observed in zebrafish embryos by Erez Raz (Fig. 5.17, right column and Movie 5.5). Three populations are involved: (\mathcal{T}_1) a population of packed cells secreting a ligand (in red), (\mathcal{T}_2) a few migrating cells (in green) that are attracted by the ligand, and (\mathcal{T}_3) a vast majority of inactive cells (not shown in the real images; shown in gray in the simulations). The migrating cells orient their polarization axis according to the chemotactic mode and exert monopolar protruding activity over their neighbors. The parameters of the passive adhesion forces are the same for all cells in this experiment.

Experiment IV results

We observe that the simulated cells (Fig. 5.17, left column and Movie 5.4) behave in a fashion similar to the real cells imaged by Erez Raz, validating the notion that the same “active” force model based on polarization axes and protrusion can underlie various cell motility behaviors such as sorting, intercalation, and migration.

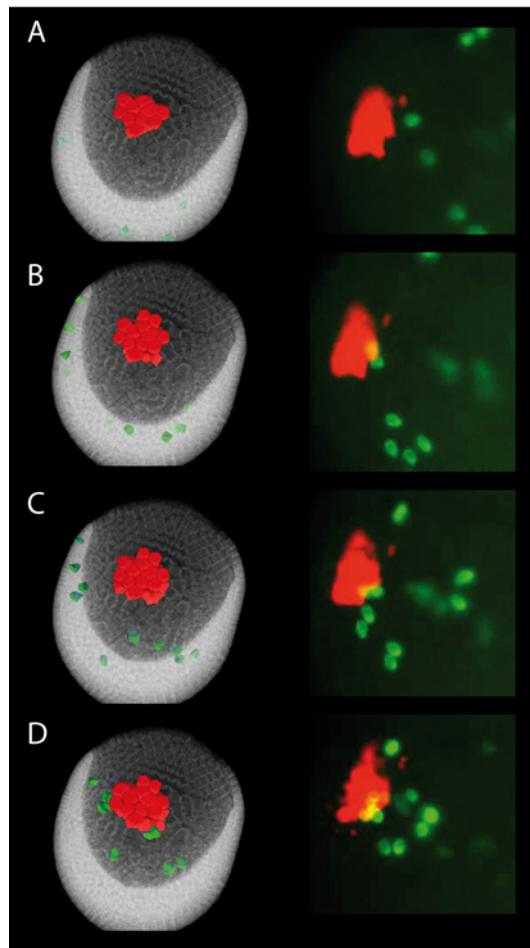


Figure 5.17: Experiment IV: individual cell migration. Left: Simulation involving a three-type cell population: a few cells secrete a ligand (the red color represents its thresholded concentration), the green cells are exerting monopolar protrusion over all their neighbors oriented uphill along the gradient, and the gray cells are inactive. Right: Microscope images of germ cells (green) migrating toward central cells expressing SDF-1a (red), which serves as a guidance cue. Images from the right column adapted from the E. Raz Lab's website <http://zmbe.uni-muenster.de/institutes/izb/izbres.htm>

Part II

STUDY OF THE ZEBRAFISH EARLY EMBRYOGENESIS

CHAPTER 6

Review of the Zebrafish Early Developmental Mechanics

Our goal in this chapter is to identify the different phases of the developing early zebrafish embryo that our model should account for, and the different components with their characteristic scales that are expected to be at play and underlie the biomechanics of this process.

The zebrafish gastrulation, its processes and underlying causalities have long been and remain today a very active field of research. The global view that we have formed about this phenomenon, and which we present in this chapter, is based on raw microscope observations, the 3D+time imaging and reconstructions performed by the BioEmergences platform, and the state-of-the-art literature surveyed in the following publications:

- D'Amico, L.A. & Cooper, M.S., 2001. Morphogenetic domains in the yolk syncytial layer of axiating zebrafish embryos. *Developmental dynamics: an official publication of the American Association of Anatomists*, 222(4), pp.611–624. [247]
- Montero, J.-A. & Heisenberg, C.-P., 2004. Gastrulation dynamics: cells move into focus. *Trends in cell biology*, 14(11), pp.620–627. [248]
- Solnica-Krezel, L., 2005. Conserved Patterns of Cell Movements during Vertebrate Gastrulation. *Current Biology*, 15(6), pp.R213–R228. [249]
- Solnica-Krezel, L., 2006. Gastrulation in zebrafish – all just about adhesion? *Current opinion in genetics & development*, 16(4), pp.433–441. [250]
- Rohde, L.A. & Heisenberg, C.-P., 2007. Zebrafish gastrulation: cell movements, signals, and mechanisms. *International review of cytology*, 261, pp.159–192. [251]
- Heisenberg, C.-P. & Solnica-Krezel, L., 2008. Back and forth between cell fate specification and movement during vertebrate gastrulation. *Current opinion in genetics & development*, 18(4), pp.311–316. [252]
- Chan, T.-M. et al., 2009. Developmental gene regulatory networks in the zebrafish embryo. *Biochimica et biophysica acta*, 1789(4), pp.279–298. [44]

- Blanchard, G.B. & Adams, R.J., 2011. Measuring the multi-scale integration of mechanical forces during morphogenesis. *Current opinion in genetics & development*, pp.1–11.[164]

We interpret the early zebrafish embryo biomechanics to result from the interaction of the blastoderm cells isolated from the yolk cell by the 32-cell stage (or somewhat later), which first differentiate in the epithelial enveloping layer (*EVL*) and the deep cells, with the *deep cells* giving rise to the *hypoblastic* and *epiblastic* layers through gastrulation (Fig. 6.1). The epiblast is described as fated to neural and non-neural ectoderm and the hypoblast is fated to the mesoderm and the endoderm.

Cell biomechanics is expected to result from the dynamics of the cytoskeleton. The tissue biomechanics should result from cell dynamics, including adhesive properties and intrinsic motility, cell-cell and cell-extra cellular matrix (*ECM*) interactions. In addition, the yolk contribution to the embryo biomechanics during early embryonic stages is highly debated [103].

It is one of our main modeling challenges to identify and integrate the components that we hypothesize to be causal in the zebrafish early embryo morphogenesis. During our more detailed exploration of the possible ingredients of the model, we will follow the known chronology of events, keeping in mind at the same time that the classical narrative of the developmental stages might also be hiding other changes in the cell dynamics. It should be one of the strengths of the model to point to the transitions in the system's dynamics.

6.1 Cleavage Stages, Formation of the EVL and YSL

We call *cleavage stages* the steps going from the 1-cell stage to the 1000-cell stage through successive cell divisions. As emphasized by Olivier et al. [254], cell cycles gradually desynchronize from cell cycle 2. The first 10 cell cycles leave only time for a succession of S and M phases, and the observed lengthening is interpreted to reflect the exhaustion of maternal components required for the underlying biochemistry. Major aspects of these steps are, in addition, the absence of cell intrinsic motility, the acquisition of cell diversity in terms of cell environment and cell volume—although the global cell volume is supposed to remain constant with the sum of the daughter cells volume equal to their mother's volume without growth. The latter is actually expected to hold true until the end of gastrulation, i.e. until cycle 15 or so. Shaping the zebrafish blastoderm at these early stages is also based on cell division orientation, and precise quantitative data has been obtained providing rules of orientation from one division to the other and relative to the embryo surface [254]. Two major morphogenetic events should additionally be singled out: the formation of the enveloping layer (EVL) and the formation of the yolk syncytial layer (YSL).

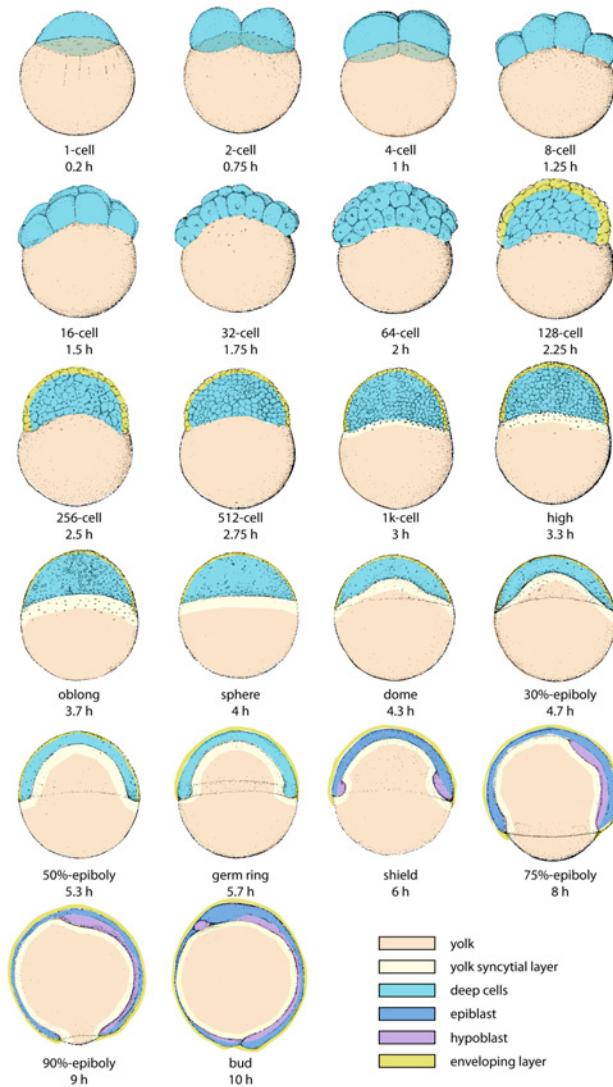


Figure 6.1: The zebrafish early developmental stages.
 Adapted from Kimmel [253]. From the 1-cell stage until the end of gastrulation, lateral views with animal pole to the top and dorsal side, identified by the shield stage, to the right. The enveloping layer (EVL) is in yellow. The deep cells (in blue) lead through gastrulation to the hypoblast (in purple) and the epiblast (in dark blue). The whole spatio-temporal sequence is expected to last 10 hours at 28.5 degree celsius.

6.1.1 The Formation of the Enveloping Cell Layer (EVL)

The EVL, long considered as an extraembryonic tissue, might rather be homologous to mammals' epidermis as recently pointed by the work of D. Wagner and coll. (Communication, Madison zebrafish meeting 2012). When formed and anchored on the yolk membrane, the EVL will bring a major biomechanical constraint to the blastoderm and to the whole embryo.

EVL differentiation is described to start as early as the 64-cell stage, and it should be a compartment by the sphere stage (4 hpf). This means that at some point inbetween these two stages, the outer layer of the blastoderm will be characterized by specific dynamics and biomechanical properties distinct from the blastoderm deep cells. Indeed, EVL is revealed as the external layer of the deep cells during the mid-blastula transition. Its cells start flattening along the radial axis and their cell cycle lengthens [255].

According to Kimmel [256], EVL cell division produces deep cells as late as the 9th cell cycle. It is shown by injections of lineage tracer dye into single blastomeres in midblastula embryos, which yield clones contributing cells to several tissues. In the example shown in Fig. 2 of [256], "a cell of the surface enveloping layer (EVL) of the blastoderm was injected at the 1000-cell stage (3 hpf); and eventually gave rise to three classes of descendants at three separate locations: ... EVL, two adjacent somites, neural cells in the spinal cord". By the sphere stage (4 hpf), EVL cell division orientation is confined to the tangential plane of the blastoderm [257]. However, mechanisms underlying the EVL differentiation are still debated [86][258].

6.1.2 The Formation of the Yolk Syncytial Layer (YSL)

By cycle 9 (512-cell stage) starts the "mid-blastula transition" (MBT) and the zygotic transcription begins. At the same time, the yolk becomes a syncytium. The yolk is described at first as a large sac filled with yolk platelets and fat droplets, which are surrounded by a lipid bilayer membrane and structured by an actomyosin cortex organized soon after the egg is laid down. The yolk becomes a syncytium through one round of cell divisions at the blastoderm margin, fusing half of the mother cells' progeny with the yolk. We then distinguish the yolk syncytial layer (*YSL*) where the yolk syncytial nuclei (*YSN*) lie, from the yolk cytoplasmic layer (*YCL*). The YSL is further divided into the internal and the external YSL. The YSL becomes an active region of transcriptional activity contributing to the dorso-ventral symmetry breaking [259]. The MBT certainly marks major changes in terms of the system's biomechanics. Cells acquire an intrinsic motility, the EVL differentiates and the yolk syncytium cytoskeleton further organizes.

6.2 Blastula Stages and the Onset of Epiboly

The causes of the blastoderm shape changes after MBT (3 hpf) for the next hour, i.e. until the sphere stage and the first manifestations of epiboly, remain elusive. The qualitative description highlighted by the actual naming of the developmental stages: “high”, “oblong”, and “sphere”, points to the flattening of the blastoderm and suggests that the overall shape of the embryo gets closer to perfectly spherical. An increase of cell-cell tension compared to yolk-cell tension might lead to the flattening of the blastoderm. Other issues might include the EVL differentiation and modification of the yolk cytoskeleton organization and dynamics.

By the sphere stage (4 hpf at 28.5 degree Celsius), the YSL and YCL are well organized with dense networks of microtubules [260][261]. Two different microtubule networks are found in the yolk: one in the anuclear yolk cortical layer, which is oriented along the animal-vegetal axis toward the vegetal pole; the other in the yolk syncytial layer, linking YSN mitotic and interphase microtubules. No microtubules were detected in the deeper, yolk-containing center of the yolk cell, but it might be an experimental artifact [262] (Fig. 6.2).

The dynamics of the yolk cytoskeleton including microtubules-based trafficking and tensegrity, and actomyosin-based tension, contraction and elastic properties, should correlate with morphogenetic events and various kinds of signaling including calcium fluxes. A recent publication summarizes this latter issue (Fig. 6.3). The schematic view below (Fig. 6.3) suggests specific features for the transition during the blastula period corresponding to the first phase of epiboly. We effectively have to account for the so-called doming phase and the respective contribution to this morphogenetic step of the blastoderm EVL tension, deep cells movements, and possibly the yolk cortex tension’s relative weakness at the animal pole.

EVL cells are obviously stretched, and as Rohde (2007) [251] puts it: “It is possible this increase in surface area is not only a passive response to EVL stretching, but also an active component of epiboly. Experiments in *Fundulus* showing an increase in apical membrane turnover in EVL cells under tension support this idea.” This view is also supported by the work of Fink et al. [264].

As shown from lineage tracing experiments [256], the progeny of injected deep cell starts to loose their cohesion during the 4th-hour period of development, indicating their newly acquired intrinsic motility [265]. From the sphere stage to the shield stage, the deep cells undergo radial intercalation. How much of this feature is the cause or consequence of this first phase of epiboly is unclear. Radial intercalation might indeed entrain the blastoderm margin and lead the epiboly movement of both the EVL and YSN. Deep cells might also actively migrate on the yolk and on the EVL to produce a similar phenotypic feature.

The role of the yolk in this first phase of epiboly might be envisioned at different levels. We already suggested a possible relative weakness of the yolk cortex that would contribute to the doming initiation. Then, although the literature does not

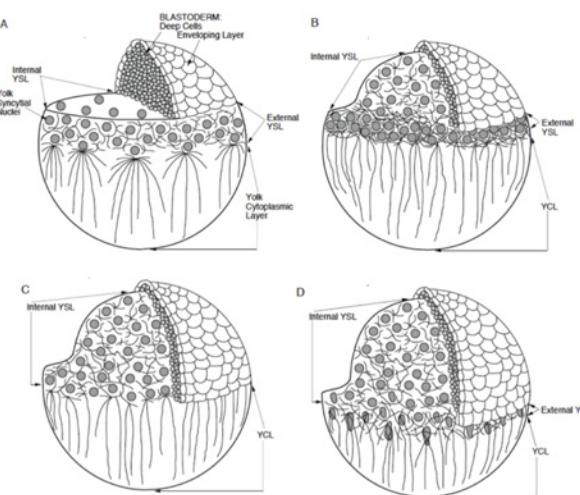


Figure 6.2: Schematic illustration of the changes in the organization of the cortical cytoplasm of the yolk cell in relation to other cell types in zebrafish embryo during epiboly in normal embryos. The organization of microtubule networks in the YSL and YCL observed in this study is indicated by thin lines. Only part of the blastoderm is shown to reveal the morphology of the yolk cell. The relative sizes of elements are not proportional. (A) The late blastula just before the onset of epiboly (sphere stage, 4.0 h). The blastoderm, composed of the internal deep cells and the superficial enveloping layer (EVL), is positioned atop of the syncytial yolk cell. The animal surface of the yolk cell underlying the blastoderm is flat. Most of the yolk syncytial nuclei (YSN) are in the external yolk syncytial layer (external YSL) vegetal to the blastoderm. The microtubules of the external YSL form a network. The organization of microtubules in the internal YSL at this stage of development is not known. The microtubules of the a nuclear yolk cytoplasmic layer (YCL) radiate from the organizing centers associated with the vegetal-most YSN and are aligned along the animal-vegetal axis. (B) 30 percent epiboly (4.7h). The blastoderm covers 30 percent of the yolk cell that bulged toward the animal pole taking on a dome shape. The external YSL has contracted and exhibits densely packed YSN and a dense network of microtubules. The external YSL is partially covered by the expanding vegetally blastoderm. (C) 50 percent epiboly (5.2h). The blastoderm arrives at 50 percent of the yolk cell latitude and covers almost completely the YSN which is also migrating vegetally and the YSL microtubule network. Only the YCL with its array of the animal-vegetal microtubules is visible vegetally to the blastoderm. (D) 60 percent epiboly (6.5h). Deep cells cover 60 percent of the yolk cell. The YSN nuclei are now visible in front of the blastoderm and lead the epibolic movement. The YSN are often stretched along the animal vegetal axis. The EVL rim is closer to the vegetal pole than the margin of deep cells. The YCL is diminished. Figure and caption from Solnica-Krezel and Driever (1994) [262]

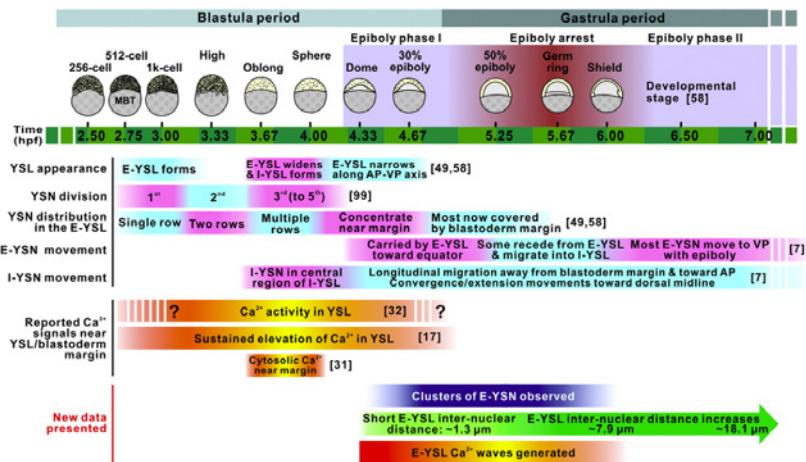


Figure 6.3: Characterization of Ca(2+) signaling in the external yolk syncytial layer during the late blastula and early gastrula periods of zebrafish development. Timeline to show the major events that occur during: the appearance of the YSL; the division, distribution and movement of the E-YSN; the movement of the I-YSN; and the reported Ca2+ signals generated at or near to the blastoderm margin, during the blastula and early gastrula periods. New data regarding E-YSN clustering, changes in inter-nuclear distance in the E-YSL over time and the generation of E-YSL Ca2+ waves are also shown. MBT, midblastula transition; AP, animal pole; VP, vegetal pole. These features correlating events at different scales does not bring us closer however to an integrated model of zebrafish blastulation and gastrulation. But it indeed provides some hints in the ingredients that the model should take into account. Figure and caption from Yuen et al., 2012 [263]

mention this possibility, we might think about an active role of the yolk content, platelets and lipid drops, if moved by an internal cytoskeleton and molecular motors. To our knowledge, however, none of these are documented in the literature. Certain functions of the yolk cortical layer and its cytoskeleton are documented. Microtubules in the yolk cortical layer (iYSL) are absent in nocodazole treated embryos [262]. Consequently, the yolk cell acquires a more spherical shape and in addition, “the YSN are blocked in their movement towards the vegetal pole and do not exhibit elongated shapes. Deep cells move very slowly toward the vegetal pole and almost cover the YSN. Epiboly of the EVL is slower than in control embryos.” However, epiboly is not impaired in nocodazole treated embryos and the process robustness indicates the synergistic involvement of multiple factors. Actomyosin contraction at the margin has been involved in the second phase of epiboly, although recent studies dispute its role [103]. Phase 1 epiboly would be however rather counteracted by a contraction at the blastoderm margin, with eventually the blastoderm slipping away from the yolk.

6.3 Gastrulation Movements

Gastrulation *per se* starts by 6 hpf and is characterized by the involution of the endomesoderm most prominent on the dorsal side, the second phase of epiboly, and the beginning of convergence-extension movements. During gastrulation, it looks like blastoderm cells, YSN, and probably also the extra cellular matrix (ECM) have the same kind of movements. As a rough approximation, all cells and tissues move in a similar fashion. They spread over the yolk toward the vegetal pole (epiboly) and converge toward the dorsal side of the embryo to form the antero-posterior axis (convergence-extension).

However, the detailed observation of cells and tissues kinematics reveals that each tissue has specificities and their mechanical contribution to the gastrulation movements may be different. This issue is of importance to implement a relevant model of gastrulation processes.

Whether cells are active or passive, with homogenous behaviors within a tissue, exerting or not forces on neighboring tissues is difficult to assess. Perturbations, whether genetic, pharmacologic, or mechanic, have been extensively used to discriminate between different scenarios.

6.3.1 Epiboly Phase Two

Although epiboly might be considered a continuous process, it probably relies on different processes distinguishing Phase 1 from Phase 2. Phase 2 epiboly coincides with the onset of gastrulation *per se*. Again, gastrulation and epiboly seem robust to a number of perturbations and it seems possible to broadly decorrelate

different processes contributing to normal gastrulation: epiboly, hypoblast formation/internalization, and convergence-extension.

What the driving force of epiboly is, however, has not yet received a definitive answer. Several hypotheses have been debated and it is not clear whether any of them should be definitely included or excluded. In fact, the epiboly process seems so robust that it very likely combines several of these hypotheses, which might even compensate for each other and synergize to produce the full-fledged phenomenon. Finally, epiboly is described for the YSN, deep cells, and EVL, and the three tissues seem to be able to proceed in part independently. We review here some of the mechanisms thought to underlie the vegetal-oriented displacement of the EVL, deep cells, and eYSL margin:

- migration of EVL cells on the yolk
- active intercalation in the EVL or active EVL surface increase
- microtubule-driven vegetal-ward movement of YSN
- endocytosis in the external YSL
- active contraction of the actomyosin ring-like structure in the eYSL margin and EVL margin
- friction force induced by a retrograde flow in the eYSL
- continuous internalization treadmill of deep cells

Possible roles of the YSL in driving epiboly

The first explanation of epiboly in teleosts was proposed by J.P. Trinkaus in 1951 [266]. He ruled out suppositions that epiboly was due to differential growth mechanisms or to a contractile activity of the yolk surface (the “yolk gel layer” idea). Instead, he favored the hypothesis that the so-called periblast (now called the internal yolk syncytial layer iYSL) is responsible for the cells’ *mass movements*. This claim mainly relied on microdissection experiments in which the blastoderm was removed, leaving the periblast exposed. In such cases, the periblast still completed epiboly until the blastopore closure. Later studies [267][268] confirmed this insight and designated the YSL as the driving force of the Fundulus epiboly. Several features of the YSL dynamics have been investigated for their possible involvement in driving epiboly, including (i) endocytosis in the eYSL, and (ii) microtubule-driven YSN movement:

- (i) Endocytosis has been observed in Fundulus just below the advancing margin of the blastoderm [269], thus in the eYSL only. In the zebrafish, the endocytotic process was observed even in the absence of YSL microtubules [262]. It is however unlikely that they play an active role in the epiboly of

the margin, as the endocytic vesicles integrate the surface of the eYSL more vegetally and do not integrate the surface of the iYSL, thus not contributing to its expansion [270].

- (ii) Even if microtubules are disrupted by nocodazole treatment, deep cells and EVL still performs partial epiboly whereas YSN stay in the 40 percent region. Thus YSN movement toward the vegetal pole is not a candidate to drive tissue epiboly [262].

Possible roles of the EVL in driving epiboly

The possibility that the EVL either migrates over the yolk surface, or rearranges, or contracts, and by doing so, contributes in different ways to epiboly, has also been discussed. Filopodial activity has been observed on the apicolateral surfaces of the EVL cells at the EVL/eYSL interface. These microspike-like structures suggest that EVL may have an active behavior at the margin [271]. However, a strong cohesion of the yolk membrane and EVL margin, with the presence of tight junctions seems to exclude the migration of EVL cells on the yolk membrane [269]. Keller and Trinkaus observed cell rearrangements in the enveloping layer suggesting that the EVL may be an active component of epiboly [272]. EVL cells, however, do not perform large-scale rearrangements during epiboly. In particular, nothing has been described so far like the “rosette” scenario in *Drosophila* [273]. We interpret that epithelia that perform active rearrangements may have a lateral-to-apico/basal length ratio much closer to 1 than is observed in the EVL. Finally, an actomyosin contraction in the YSL induces EVL cell shape changes [274], correlating with margin contraction, and the resulting constriction has long been postulated to be the driving force of epiboly.

The purse-string scenario, contraction at the margin of the eYSL and/or EVL margin

The presence of F-actin-based structures at the margin after 50 percent epiboly was described as follows by Cheng [275]:

“They are composed of two ring-like F-actin structures that form at the deep cell and enveloping layer margins of the blastoderm and a punctate actin band that develops in the external yolk syncytial layer. Treatment with cytochalasin B or the calcium chelator dibromo-BAPTA results in the disruption of all three of these actin-based structures, leading to the slowing or immediate arrest of epiboly, respectively, followed by a failure of yolk cell occlusion and the eventual lysis of the embryo through the vegetal pole region.”

These observations suggest the role of a contractile ring in the occlusion of the vegetal portion of the yolk cell during the Phase 2 of epiboly [275]. A hypothesis is

that the Phase 2 epiboly is driven by an actomyosin contraction at the margin of the EVL. The local density of actin at the EVL margin correlates with its deformation. And the process involves actin and myosin 2 recruitment within the yolk cytoplasm, as observed in *Drosophila* for the dorsal closure [274]. This scenario has been recently challenged and the state of the art in the domain now comes from the following works, described next.

A friction force induced by a retrograde flow in the eYSL

A recent study hypothesized the existence of an additional and previously undescribed force that would pull the EVL-yolk margin vegetally by a flow-friction mechanism [103]. This mechanism is shown to be sufficient to drive epithelial epiboly after 50 percent epiboly because, in mechanically constrained cylindrical embryos where a purse-string scenario is inefficient, the margin is still converging to the vegetal pole. We propose however that radial intercalation in the epiblast could explain the sustained epiboly in cylindrical embryos. This hypothesis might be tested by *in vivo* imaging and assess the position of the blastoderm margin compared to the EVL margin.

Possible role of deep cells dynamics in driving late epiboly

Epiboly might emerge from the behavior of deep cells. Relevant behaviors include cell intercalation, radial and/or medio-lateral through convergence and extension, and cell involution at the margin. However, Trinkaus's work in *Fundulus* led to the idea that deep cells are not responsible for the late epiboly ("Epiboly of *Fundulus*, and presumably of teleosts generally, does not depend on the deep cells." [268]). Indeed, colchicine treatment of the *Fundulus* gastrulae, blocking microtubules polymerization, impaired the deep cells mechanical properties, and yet the epiboly still performed normally [276]. Moreover, the possibility that hypoblast involution contributes to epiboly was excluded by authors interpreting that involution is not a continuous process throughout epiboly [277]. Our observations suggest that this conclusion might be revised (e.g. Movie 6.1). We propose that a continuous internalization treadmill of hypoblastic cells might both drive deep cells epiboly and entrain the EVL.

In addition, other work mentions deep cells' intercalation as contributing to epiboly. Epiblast deep cells intercalate radially from the interior layer to the sub-EVL layer. They change their shape to actively drive the epiboly. Compromised E-cadherin function exhibit reduced radial cell intercalation movements [278]. An E-cadherin mutant, "half-baked" (*hab*), expressed in a radial gradient in the nonaxial epiblast, induce cell phenotype that do no maintain their intercalated position, and "de-intercalate" toward their previous interior layer [279]. In addition, *hab* mutation stops margin epiboly but it hardly affects internalization and convergence extension.

Moreover EVL and YSN produce unaffected epibolic movement, so they do not need proper deep cell epiboly [280].

6.3.2 Hypoblast Formation

The relative movements of the YSL, deep cells, and EVL leave open the possibility that one is entraining the others. D'Amico and Cooper showed that the YSN and blastoderm cells' movements during gastrulation are similar [247]. C.P. Heisenberg excluded however the possibility that the YSNs pulled the hypoblast, by showing that they move behind the hypoblast anterior border [281]. A more systematic quantitative study would certainly be useful to revisit these issues, based on *in vivo* imaging, such as shown in Movies 6.2 (from [282]) and 6.3 (from [248]), which prompt a number of questions about the dynamics of hypoblast involution, the relative movements of deep cells and EVL, or the cells' neighborhood rearrangements at the onset of gastrulation.

The general description of germ layers formation and gastrulation movement is still qualitative and remains vague in terms of the individual cell behaviors underlying morphogenetic events, as stated in Montero et al. [278]:

“In zebrafish, the first mesendodermal progenitors are induced at the margin of the blastoderm when the blastoderm starts to spread over the yolk cell (dome stage) (for reviews, see Kimelman and Schier, 2002; Warga and Stainier, 2002). When the blastoderm covers about half of the yolk cell (50 percent epiboly), the germ ring forms as a local thickening at the margin of the blastoderm. Germ ring formation is accompanied by convergence movements of blastodermal cells, leading to a compaction of cells at the dorsal side of the germ ring, where the embryonic organizer or ‘shield’ forms (Warga and Kane, 2003). This is also the time, when mesendodermal progenitors within the germ ring begin to internalize by moving first to the margin of the blastoderm, then downwards in direction of the yolk sac and eventually migrating back towards the animal pole of the gastrula (Warga and Kimmel, 1990).”

Ingression and involution

The cell behaviors underlying the formation of the hypoblast (mesendodermal layer) are described as a combination of ingression and involution. It is however not clear that the concept of *involution* described by Trinkaus [283] applies to the gastrulating zebrafish hypoblast:

“*Involution* is the flowing of a sheet of cells over the edge of an in-pocketing, where invagination has occurred, such as at the blastopore during amphibian gastrulation.”

Trinkaus believes that *Fundulus* does not perform a “classical involution” as observed in *Amphibia*. A cohesive sheet of cells is supposed to internalize by involution in opposition to *ingression* for “individuals or small clusters of individuals” that “sink inside as individuals, changing their relations to each other in the process”. The motility pattern is different. In the involuting tissue, there is no internal topological rearrangement between the cells. Ingressing cells have extensive topological rearrangements and may even be mixed with the non-internalizing tissue.

In the case of the zebrafish gastrulation, we use the involution concept as an archetypal description characterizing the motion of a collective sheet of cell plunging toward the yolk mass and turning back toward the animal pole. But we interpret that cell behaviors leading to the hypoblast formation, their specific features depending on the position along the dorso ventral axis, and the processes that segregates the hypoblast and the epiblast, all remain open questions. We can indeed find studies concluding about hypoblast ingression and others rather suggesting involution. D’Amico and Cooper [247] suggest that ingression is not the dominant mode of internalization. This is the interpretation derived from 4D confocal imaging of 114 fluorescent-dextran labeled cells internalizing at the dorsal marginal zone. “98 percent of the cells entered the hypoblast within approximately three cell diameters from the edge of the margin”. 70 over 114 cells entered the hypoblast within a 1-cell diameter from the margin. This argument of the distance from the margin where cells enter the hypoblast does not so clearly favor the internalization at least on the dorsal side for the prechordal plate population in the shield. Montero et al. [278] observed that:

“Notably, all of the prechordal plate progenitor cells – out of the 70 epiblast cells we analyzed in dorsal/axial of the germ ring – delaminated not more than 4-5 cell diameters away from the tip of the germ ring; this indicates that prechordal plate progenitor cell internalization is restricted to the marginal-most region of the germ ring.”

***MZoep* mutants**

The investigation of hypoblast formation defects in *MZoep* mutants devoid of both the maternal and zygotic components of the Nodal pathway indicate that hypoblast cells rather ingress and point to the cell autonomy of the process [284]. The ingression hypothesis is also favored in Montero, 2005 [278]: internalizing cells delaminate as single cells within more than “4-5 cell diameters away from the tip of the germ ring”.

In *MZoep* mutants internalization does not happen at all. *MZoep* mutant cells transplanted to the margin of wild-type embryos internalize normally but do not migrate toward the animal pole. In the reverse transplantation experiment WT transplanted cells internalize properly and express mesendodermal markers.

"Interestingly, in *MZoep* embryos the wild-type cells internalized by directly moving deep toward the yolk without first moving to the most vegetal region of the margin, as in wild-type embryos."

The mechanisms of this individual ingression are however unknown. We interpret that this observation does not necessarily imply an intrinsic motility of the cells. The WT transplanted cell diving to the yolk might be excluded by the *MZoep* host cells as they intercalate. Such a hypothesis might be tested in the context of *cadh1* loss of function where impaired radial intercalation might perturb the process. The behavior of WT cells in *MZoep* mutants is reminiscent of the behavior of cells with constitutive activity of the nodal pathway when transplanted in a wild type background [285]. This suggests that the major issue might be some differential between the donor and the host in terms of adhesion, polarity, intrinsic motility. A quantitative assessment of these features might help highlighting this point.

Timing and separation

In addition to the mode of internalization: involution or ingression, the timing of involution, meaning onset and duration of the process are not so precisely documented. The onset of internalization might be indicated by the formation of the germ ring that becomes visible almost simultaneously all around the embryo, within 15min [265]. The duration of internalization is also an issue and quantitative data is largely lacking. According to Keller et al. [277], involution is achieved through an embolic phase. Our observations from *in vivo* imaging suggest however that involution might go on throughout the whole gastrulation. This hypothesis is very likely to have consequences on the dynamics of epiboly that we should be able to test with our modeling strategy.

Hypoblast involution means that at some point, the epiblast and the hypoblast are distinct compartments. Again, the modalities of this separation should be revealed by the cell lineage data. If the relative movements of the two cell populations far from the margin suggest that there is probably no passage from one to the other far from the blastoderm margin, segregation close to the margin depending on the involution/ingression movements might lead to a fuzzy border. In addition, hypoblast needs to be separated in endoderm and mesoderm. The former is described to involute first, have a cell autonomous fate, and spread over the yolk through a random walk [285].

Later progression

How cells behave after internalization/ingression to progress throughout epiboly is another issue. Three distinct populations with relatively autonomous behaviors progress both anteriorly toward the animal pole and posteriorly to reach the tail bud: the endoderm and the mesoderm that form the hypoblast, and the ectoderm.

The endoderm is a mono-layer of highly motile and rather sparse mesenchymal cell (Movie 6.4 from [286]). We hypothesize that the kind of contact inhibition observed in Drosophila cells might explain the dispersed pattern of endodermal cells [287]. The mesoderm becomes regionalized in sub-populations with distinct behaviours and fates: the prechorale plate and the notochord at the dorsal midline, the paraxial mesoderm, lateral and ventral mesoderm. The mesendodermal layer has been described to migrate collectively with and intrinsic directionality given through cell contacts¹. The ectoderm is the neural ectoderm dorsally and non-neural ectoderm ventrally. The refinement of tissues regionalization with distinct fates should early on correlate with distinct cell behaviors that have not been quantified so far.

6.3.3 Convergence-Extension

As gastrulation proceeds, movements of convergence toward the dorsal side and extension of the tissues along the animal/vegetal becoming the anterior/posterior axis increase and culminate around the tailbud closure. Convergence-extension leads to a marked dorso-ventral asymmetry that might have been initiated early on by the maternal distribution of the Nodal ligand Sqt, reinforced by the asymmetric distribution of nuclear β -catenine at the onset of zygotic activation and then morphologically obvious with the formation of the embryonic shield [50].

Convergence extension movements seem to be largely independent from other gastrulation movement as suggested by the phenotype of *hab* mutants, defective in the expression of E-cadherin, where epiboly stops, however hardly affecting CE [288]. A combination of cell behaviors is hypothesized to contribute to convergence with distinct features from ventral to dorsal: individual migration, cohesive migration, intercalation [289].

In addition, although all the tissues (epiblast, mesoderm, endoderm) display convergence extension movements, they do so with their respective dynamics thus having relative movements. The movements of the YSL nuclei are particularly intriguing as they are achieved in the absence of cell boundaries. It was hypothesized that the microtubule arrays that interlink YSL nuclei are necessary to achieve the process [262]. The correlation of YSL nucleus movements and blastoderm cell movements has been extensively studied [247]. Hypoblast cells do not migrate on a passive YSL as it would send YSN in the opposite direction. YSL can be active, either be autonomously converging, or actively reacting to DC migration. Blastoderm cells and more precisely the large endodermal cells are migrating over the yolk membrane separating them from the yolk nuclei, pointing to their possible influence on each other [290]. As D'Amico and Cooper put it [247]:

“Epiboly of the blastoderm has been shown to be at least partially dependent on the normal epiboly of the YSL (Trinkaus, 1984b; Strähle

¹ Dumortier:2012jk

and Jesuthasan, 1993; Solnica-Krezel and Driever, 1994; Kane et al., 1996; Zalik et al., 1999). An alternative hypothesis is that the motility of deep cells in the blastoderm is responsible for directing the YSL nuclear movements. Is it possible that cells migrating on the surface of the YSL could somehow drag YSL nuclei with them, even though the plasma membrane of the yolk cell physically separates them? Although it is formally plausible that a bulk flow of YSL membrane and YSL cytoplasm could be generated by migrating cells collectively pulling the YSL membrane, this theoretical possibility is unsupported by numerous observations. Previous studies have shown that deep cells migrate individually, in different directions, and independently of underlying YSL nuclei (Warga and Kimmel, 1990; Solnica-Krezel and Driever, 1994; Wilson et al., 1995; Trinkaus, 1996; Warga and Nüsslein-Volhard, 1999). Moreover, it is known that the epiboly movements of the YSL nuclei in both Fundulus and zebrafish are not dependent on the overlying cells of the blastoderm (Trinkaus, 1984a,b, 1993; Kane et al., 1996). Thus, it seems unlikely that the convergentextension movements of YSL nuclei in zebrafish embryos are driven by the convergent-extension movements of deep cells in the overlying blastoderm. A third possibility is that, despite their similarities, the collective movements of deep cells and YSL nuclei are largely autonomous of each other.”

In addition to convergence extension movements, oriented cell division, regulated by the planar cell polarity pathway, was hypothesized to be a driving force for axis elongation [291][292]. We expect our modeling strategies to help assess the respective contribution of protrusive activity and cell division orientation to the elongation process.

CHAPTER 7

Experimental Reconstruction of the Zebrafish Early Development

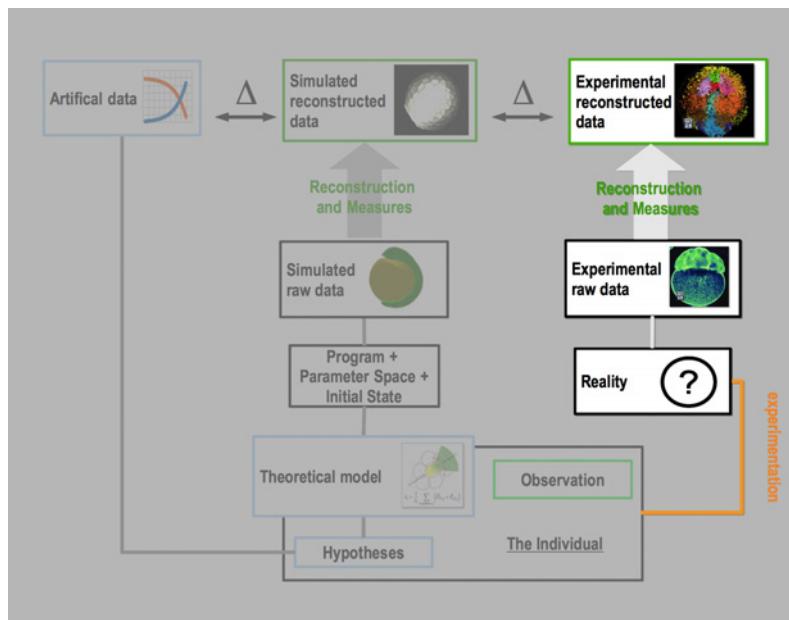


Figure 7.1: Situation of Chapter 7 in the methodological workflow.

Understanding the morphogenetic processes during animal embryogenesis requires the systematic investigation of cells' behavior and the reconstruction of the cell lineage tree as a branching process in space and time annotated with relevant measurements at the single-cell level (Fig. 7.1). The automated reconstruction of the Nematode cell lineage from confocal images established the first standards but did not provide reliable results beyond the 194-cell stage [293]. More recently, the zebrafish early embryo reconstruction from DSLM (digital scanned light sheet microscopy) images [277] or the reconstruction of Drosophila gastrulation movements from MLSM (multiphoton laser scanning microscopy) [294] also fell short of this goal as they did not provide clean single-cell tracking and were not suitable for cell clonal analysis.

In this chapter, we first summarize the strategy and tools developed by the

laboratories hosting this thesis work (Section 7.1) then describe our own addition to this workflow, custom-designed for the modeling needs of the MECAGEN project (Section 7.2). From both the biological and computational points of view (Fig. 7.1), our labs and their partners gave us the opportunity to compare our model with real data and benefit from shared computational power via a cluster and GPUs. The main asset that we could rely upon was the BioEmergences platform, developed on top of two European projects (Embryomics and BioEmergences, funded by the FP6-NEST program), which were devoted respectively to reconstructing the cell lineage tree during animal embryogenesis and measuring the variability between different individuals of the same species.

Today, the BioEmergences platform enables the quantitative reconstruction and analysis of early zebrafish embryogenesis, including the filtering and segmentation of raw imaging data acquired via LSM (Laser Scanning Microscopy) or DSLM (Digital Scanning Light Microscopy), and the automated detection of cell positions by cell tracking. Temporal sequences of 3D images can be directly reconstructed from 2D stacks and sent to a database, which also stores information about the animal model, the imaged period of development, and various technical aspects such as embryo staining and mounting, microscope settings, image size, spatial and temporal resolution. The outcome of the digital reconstruction (essentially cell tracking and segmentation) can be automatically stored in the BioEmergences database and visualized through the interactive visualization interface Mov-It. Mov-It also allows the user to manually correct cell positions and tracking, make annotations and guide pattern classification. Calculations are performed on a computing grid and distributed on many processors at several geographic locations, thus performing in a few days reconstructions that would otherwise require months on a single machine.

7.1 Bioemergences Reconstruction Workflow

In this section we describe the existing BioEmergences platform developed by our team. (Then, in Section 7.2., we will explain our own contribution to the experimental reconstructed data workflow.) First, in 7.1.1, we briefly review the protocol for zebrafish embryo preparation and imaging. Then, in 7.1.2, we introduce the image processing pipeline for the automated cell tracking, segmentation and validation, including the interactive Mov-It visualization platform with its user-friendly fast annotations and corrections via a relational database link. Finally, 7.1.3 summarizes BioEmergences' data analysis strategy and extraction of cell displacement vector fields.

The Bioemergences workflow (Fig. 7.2) is especially designed for processing 4D (3D+time) data from embryos engineered to highlight cell membranes and nuclei via the expression of fluorescent proteins (FPs). In the end, it provides a reconstruction equivalent to a “digital embryo” represented by a cell lineage tree annotated with

quantitative measurements of membrane and nucleus shapes.

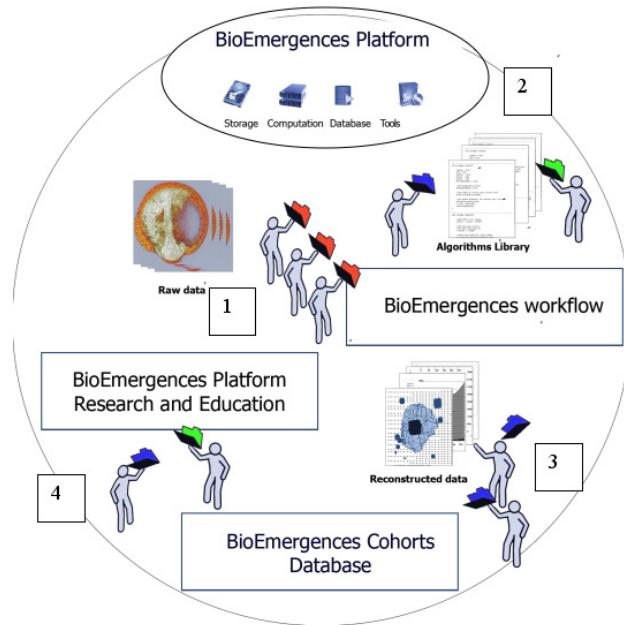


Figure 7.2: The Bioemergences platform.

7.1.1 Embryo Preparation and Acquisition

Wild-type zebrafish embryos are injected at the one-cell stage with 200pg mCherry/H2B RNA and 200pg eGFP-ras prepared from PCS2+ constructs [295][296]. Although mCherry, unlike eGFP, bleaches significantly through imaging, this color combination is the best compromise allowing proper staining of the cell membranes for further segmentation. Injected embryos are raised at 28.5C for the next 3 hours. Embryos are mounted in a 3cm Petri dish with a glass coverslip bottom, sealing a hole of 0.5mm at the Petri dish center where a Teflon tore (ALPHAnov) with a hole of 780 μ m received the dechorionated embryo. The embryo is maintained and properly oriented by infiltrating around it 0.5 percent low melting point agarose (Sigma) in embryo medium [297]. Temperature control in the room results in a temperature of about 26C under the objective slightly slowing down development with respect to the standard 28.5C developmental table [253]. After the imaging procedure, embryo morphology is checked under the dissecting binocular and the animal is raised for at least 24h to assess morphological defects. Embryo survival depends on total imaging duration, average laser power, and image acquisition frequency (or time step Δt).

Imaging is achieved with a Leica upright microscope SP5 MLSM equipped with a 20/0.95NA W dipping lens objective (Olympus) or Leica 20/1NA W dipping lens objective. Axial resolution at the sample surface (1.5 μ m) was estimated by recording

3D images of 0.1 or 1 μm fluorescent polystyrene beads (Invitrogen) at the surface of an agarose gel. Simultaneous dual wavelength excitation is performed with pulses at two different wavelengths (980 and 1030nm). At 1030 nm, pulsed laser beam (50Mhz, 200fs) is provided by a solid-state Ytterbium femtosecond oscillator (T-pulse 20, Amplitude Systèmes). At 980 nm, pulsed laser beam (80Mhz, 100fs) is provided by a Ti-Sapphire femtosecond oscillator (Mai Tai HP, Newport Spectra physics). Field size is 700×700 or 775×775 microns in x, y, 140 μm or 100 μm in z. Voxel size is 1.37×1.37×1.37 μm or 1.51×1.51×1.51 μm .

7.1.2 Cell Lineage Reconstruction Workflow

The workflow for reconstructing digital embryos is summarized in Fig. 7.3 and further details are explained below in this section. Raw data is composed of a temporal series of 3D images representing cell nuclei and membranes, which are automatically reconstructed from 2D stacks and stored in the BioEmergences database. Data is automatically sent to the EGEE computation grid through a Web interface and the output results stored in the database. Digital reconstructions can be viewed, corrected, validated and annotated through the Mov-It visualization interface.

Filtering

Data from in vivo imaging typically suffers from low contrast. This is why images are first filtered by an edge-preserving and enhancing method, the Geodesic Mean Curvature Flow (GMCF). The GMCF is able to improve the signal-to-noise ratio, faithfully preserving the position of boundaries that define the shape of the structures. Representing the input 3D image by a real function $u_0(x)$, $u_0 : \Omega \rightarrow \mathbb{R}$, where $\Omega \subset \mathbb{R}^3$ is a rectangular spatial domain, the image processing equation can be written:

$$u_t = |\nabla u| \nabla \cdot \left(g(|\nabla G_\sigma * u|) \frac{\nabla u}{|\nabla u|} \right)$$

where $u_t(x)$ is the unknown function representing the smoothed (filtered) image intensity, G_σ is a Gaussian function with a variance σ , and $g : \mathbb{R}_0^+ \rightarrow [0, 1]$ a non-increasing function representing the nuclei's and membranes' boundaries. Further details can be found in [298][299].

Cell identification

The strategy for cell identification is to assume that each cell in the dataset always contains a nucleus, and then ascribe it the position of the center of its nucleus. Nucleus centers are detected by assuming that objects in the image can be seen as “humps” of image intensity. Thus they are identified as local maxima of a further simplified version of the nucleus images, obtained either by a difference of Gaussian

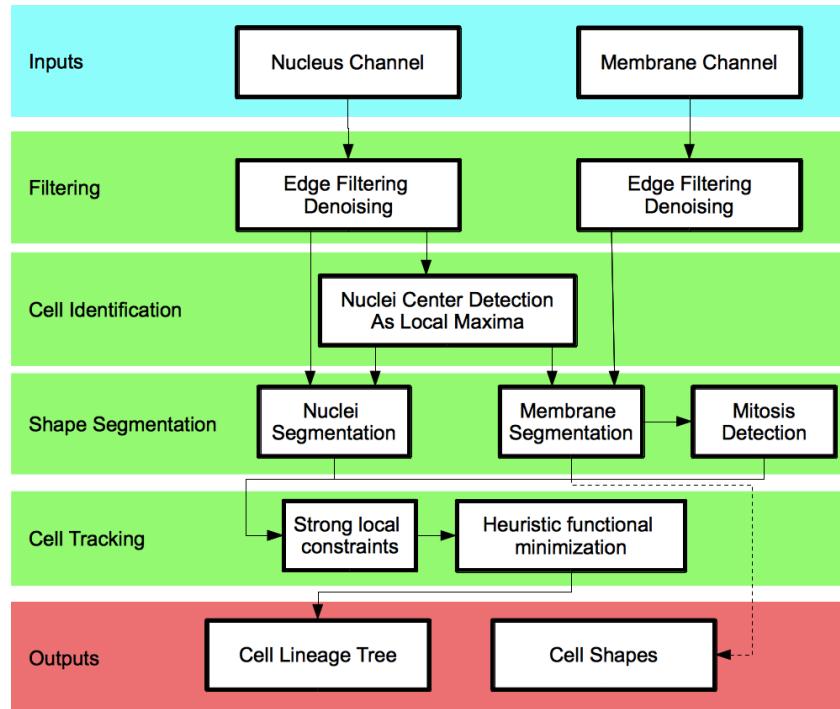


Figure 7.3: The Bioemergences workflow. Raw images of nuclei and membranes (cyan layer in the diagram) are first filtered by an edge-preserving smoothing method. Cell positions are then extracted from the local maxima of this simplified images versions, and assimilated to the nuclei's positions. They are used to initialize both the segmentation and cell tracking tasks (green layers). The final output is composed of the cell lineage tree and the nuclei/membranes segmentation (red layer).

filtering (Barbara Rizzi, oral communication) or by the nonlinear advection-diffusion equation [300][301].

Shape segmentation

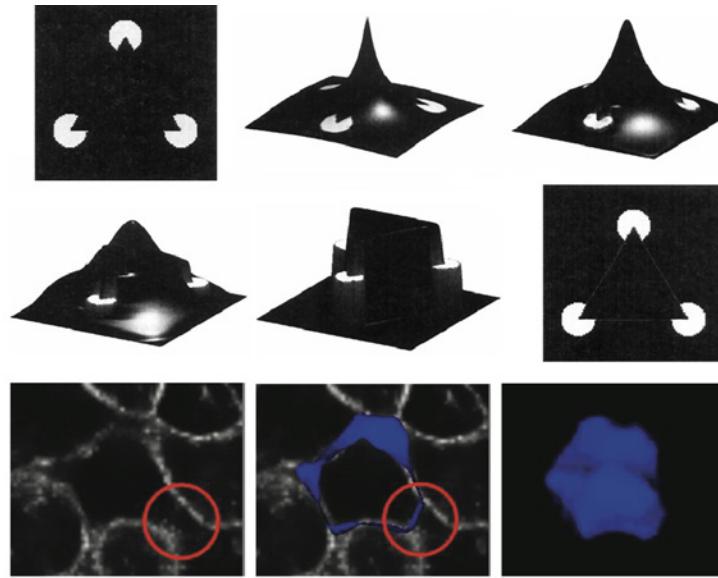


Figure 7.4: Subjective Surfaces Method. Top and middle rows: 2D example summarizing the main steps of the segmentation algorithm, illustrated on the Kanizsa triangle (from left to right): original image depicting a triangle with missing “subjective” contours [302]; construction of a surface located at the center of the triangle; evolution of the surface to minimize its area according to the model equation (two images); steady state condition giving a piecewise constant solution; choice of level set for the surface representing the contour of the triangle. Bottom: Segmentation of a membrane with missing boundary [303]. Raw data displayed as an orthoslice in the x,y plane. Incomplete membrane contour highlighted with a red circle. Segmented membrane represented as an isosurface in blue.

The segmentation of both cell nuclei and membranes is then performed on filtered data using as seeds the positions of the cell centers. Each nucleus/membrane is segmented independently on a region of interest located around the initialization point. Single segmentations are then combined to reconstruct whole volumes. The method used for the segmentation of a single object is a generalized version of the “Subjective Surface” method, whose originality is its ability to complete missing contours, making it particularly suitable for reconstructing membrane shapes. The main steps are summarized in Fig. 7.4. In the 3D case the method consists of minimizing the volume of a 3D manifold embedded in a 4D Riemannian space with

a metric constructed on the image itself. Similarly to the filtering subsection, we define here the image as a real positive function in the same image domain Ω . Initially, this function is $\Phi_0 = \Phi(x, y, z, 0)$, then it evolves over a scale parameter t and is denoted $\Phi_t = \Phi(z, y, z, t)$. With this, the model equation reads:

$$\Phi_t = gH |\nabla\Phi| + \nabla g \cdot \nabla\Phi$$

where H represents the mean curvature of the function Φ in the induced metric, and $g = g(|\nabla G_\sigma * u_0|)$, with u_0 representing the filtered image and g the image contours in a way similar to the filtering section. Further details can be found in [302][303].

Segmented nuclei and membranes are then used in the next step, cell tracking, to provide information about false positive nucleus centers and mitosis. False positive centers generally appear as two or more points inside the same nucleus and produce overlapping segmented nuclei. In this case, only one nucleus is preserved. Similarly, mitosis events are identified as overlapping segmented membranes starting from two initialization points both located inside the segmented region [304].

Cell tracking

The cell tracking algorithm is the key module of this workflow. It relates temporally the identity of a cell between two consecutive time steps. Starting with the initial cell population, these temporal links create together a tree whose branches split when a cell divides. Let I be the whole set of cells at all time steps. The unknown tracking graph f on I represents the lineage tree, and is determined in three steps:

- Initialization through an *a priori* criterion based on the nearest neighbor: the position of each cell at a given time step is linked to the position of the nearest cell in the following time step.
- Local refinement by minimizing a heuristic functional $E(f)$ whose local minima correspond to a “good” matching along cell trajectories: this functional is constructed by taking into account the heterogeneous information extracted from the processed images (maximum speed, minimum/maximum cell lifetime, probability of mitosis, changes in image intensities, vector field, cell division rate, etc.) and a biophysical regularization constraint (assumption of an elastic behavior):

$$E(f) = \sum_i \|X_{f(i)} - X_i\|_1 + \sum_i \sum_{j \in \mathcal{N}_i} \| (X_{f(i)} - X_i) - (X_{f(j)} - X_j) \|_2$$

where i and j are the indices of neighbor cells at time t , \mathcal{N}_i is cell i 's neighborhood, X_i is its feature vector, $\|\cdot\|_1$ a norm in feature space and $\|\cdot\|_2$ a norm in feature deformation space. Minimization is performed according to the simulated annealing algorithm [305].

- Global refinement by processing the obtained tree: discontinuities in trajectories are identified as false negative nuclear centers and cells that live only for a few time steps as false positive nuclear centers.

An example of cell tracking is displayed in Movie 7.1.

7.1.3 Post Processing and Exploitation

Visualization, annotation, and manual validation

After the image processing steps of the phenomenological reconstruction, the “digital embryo” can be explored and compared with the original data via the user-friendly Mov-It visualization interface. This GUI tool was designed to allow fast correction of remaining errors, recover the complete clonal history from imaged cells, and analyze cell population dynamics and behavior. Mov-It can display multiple data types in time: original images can appear either in volume rendering, orthoslices or projections; detected cell positions are represented as dots and trajectories as streamlines; segmentation results can be visualized as isosurfaces and superimposed on top of raw data with some transparency. It is also possible to easily select a cell position and observe its temporal position along the lineage tree, visualized as a set of line segments splitting into two at each mitosis event. Cell positions and trajectories can be manually corrected to validate the lineage tree. Cell populations can be classified either by automated selection based on cell statistics (such as speed, lifetime, orientation of divisions, volume, or number of neighbors) or by manual selection. For example, Movie 7.3 shows cells trajectories over time combined with a fate map until somite stage. Presumptive organs have been manually selected at the end of the acquisition, then backtracked. The net result is the creation of *fate maps*, as shown in Movie 7.2 about the interactions between prechordal plate, presumptive hypothalamus and presumptive ventral telencephalon.

Automatic validation

As manual validation by visual inspection is a very time consuming task, the BioEmergences platform also offers an automated method to validate the tracking process. The idea is to compare the tracking of the entire population of cells of the embryo (helped by manual intervention and termed “gold standard”) to the tracking of only a subpopulation in the same embryo, and assume that the accuracy of the lineage tree reconstruction in this subpopulation is as good as the gold standard. To this aim, a double labeling of cells is first obtained by transplanting about 100 cells between the high stage and the sphere stage [253] from a donor embryo with double stained nuclei (H2B/mcherry through RNA injection into H2A/eGFP transgenic embryos) into a transgenic host with green nuclei [306]. Then, tracking comparison is performed through a mapping that minimizes the distance between matching cells.

Exploitation

An example of exploitation and post-processing calculation on top of the reconstructed embryo is illustrated in the figures below (Figs. 7.5 and 7.6), which represent the cell kinematics of the embryonal structure as a velocity map and its singularity points [98].

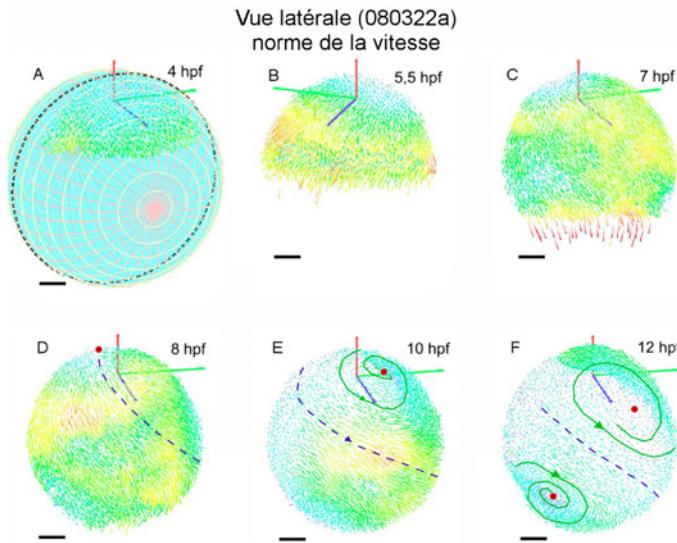


Figure 7.5: Evolution of the averaged speed field with its norm between 4 and 12 hours of development in zebrafish embryo. Dataset 080322a, lateral view. Green axis points towards the dorsal side, red axis towards the animal pole. Colormap between light blue and red related to the norm of the averaged speed field. Singular points in red, separatrices in blue, flow pathways in green. Image and caption adapted from B. Lombardot (2010) [98].

7.2 Reconstructing In Toto Deformation Fields

This section is dedicated to an additional module of the Bioemergences workflow that we developed especially for our study (Fig. 7.7). It takes in input a series of raw images from a Digital Scanning Light Microscope (DSLM) and produces in output a global “deformation field”, composed of six image processing components (explained in Sections 7.2.1-5 below).

The rationale of this additional module is the following. Biphotonic microscopy devices do not allow the capture of the entire zebrafish embryo because of optical loss of signal and image deformation in the depth of the tissue. One solution is to take advantage of the spherical shape of the embryo during early gastrulation and rotate it to capture multiple views (Fig. 7.8), thus covering the whole volume without

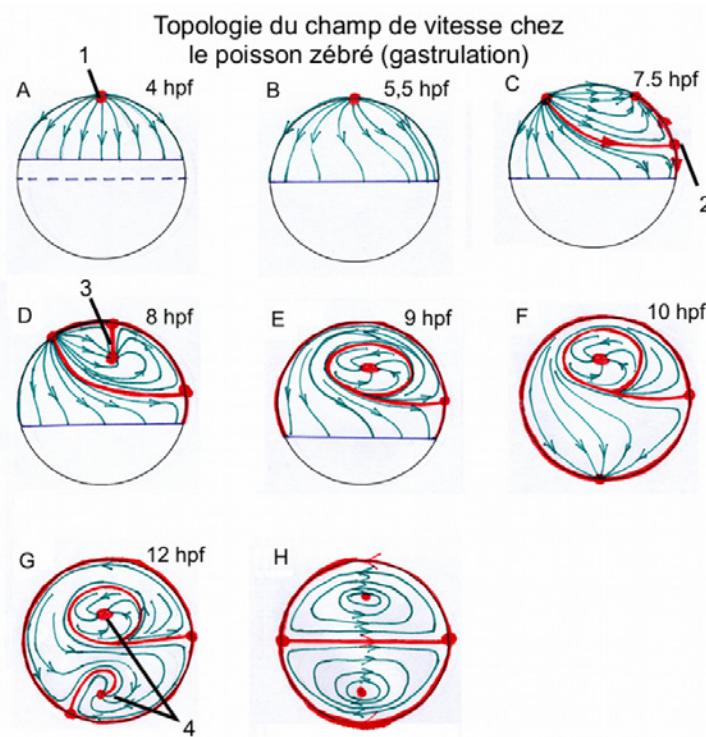


Figure 7.6: Simplified topology of the speed field in zebrafish embryo between 4 and 12 hours of development. Circles in each figure represent a schematization of the embryo in a lateral view. Dorsal side on the right, ventral on the left. Animal pole on the top, vegetal pole on the bottom. Separatrices represented as red lines, singularities as red dots. Image and caption adapted from B. Lombardot (2010) [98].

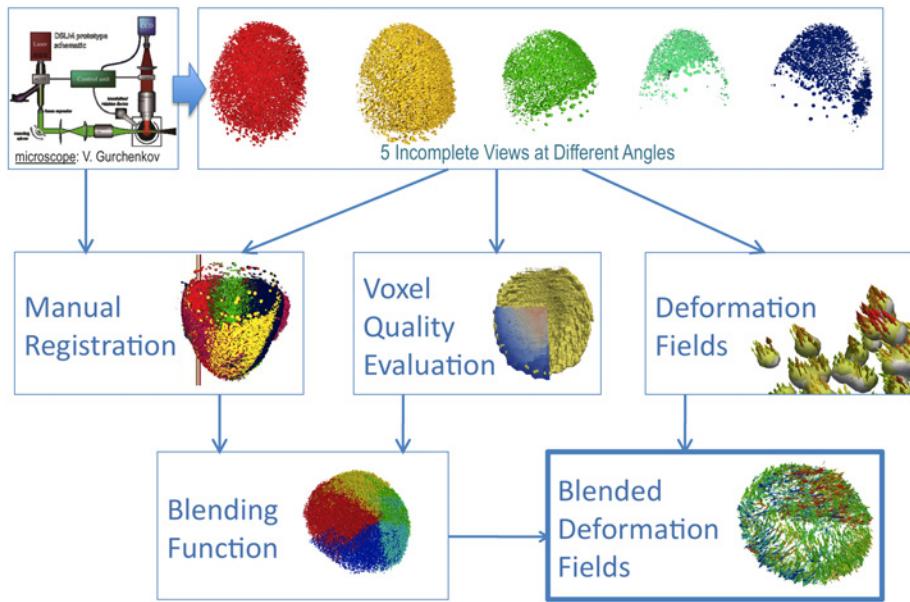


Figure 7.7: Global diagram of the reconstruction workflow of the *in toto* deformation fields.

major image degradation. The problem with this solution is that the total imaging time is greater since we need consecutive angles of view to capture superposed region of the embryo, and cells move slightly in the meantime. Fortunately, a recent microscopic technology, the Digital Scanning Light Microscope (DSLM), increases the speed of scanning thanks to its ability to illuminate and record the signal emitted by the embryo plane by plane, and not just point by point. Faster recording allows a better comparison and match of the 3D volumes to extract and reconstruct the dynamics of the cell trajectories described in the previous section.

At this point, however, we ended up with a quality of signal different from the one generated by the biphoton illuminated device used in the previous section, therefore the reconstruction strategies had to be adapted accordingly. We decided to set aside the reconstruction of the *in toto* lineage tree for later investigation, and instead work on a more accessible reconstruction: the *in toto* deformation fields of the embryo. This workflow is presented in detail in the subsections below. We programmed it using the OpenGL and ITK/VTK libraries.

7.2.1 Manual Registration

We explain here the manual registration component (Fig. 7.9). The DSLM allows rapid imaging of the embryo and, coupled with a rotating device, the recording of multiple views of the same embryo obtained from different angles (Fig. 7.10). We have been using an exact replicate of the DSLM described in [277]. In this whole

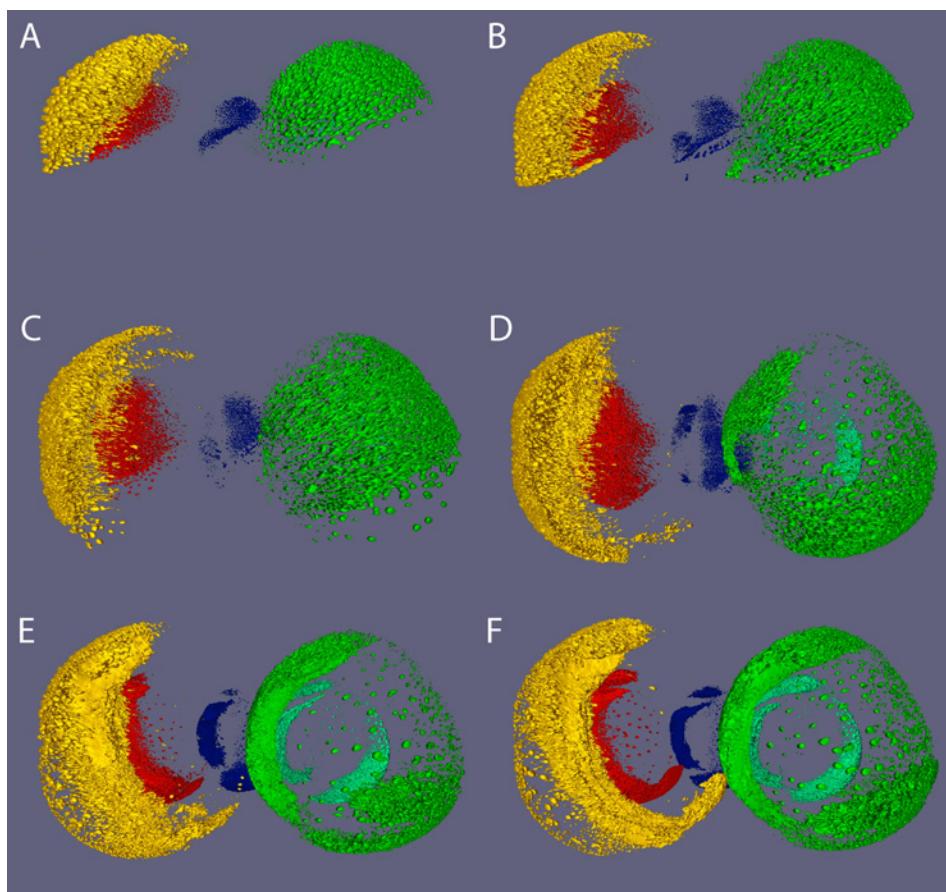


Figure 7.8: Original five views produced by the Digital Scanning Light Microscope (DSLM). Isosurfaces have been extracted from the DSLM raw data. Each view is obtained by rotating the embryo 72 degrees (one color per angle of view). A: Sphere stage. B: 40 percent epiboly. C: 75 percent epiboly. D: 90 percent epiboly. E: 8-somite stage. F: 12-somite stage.

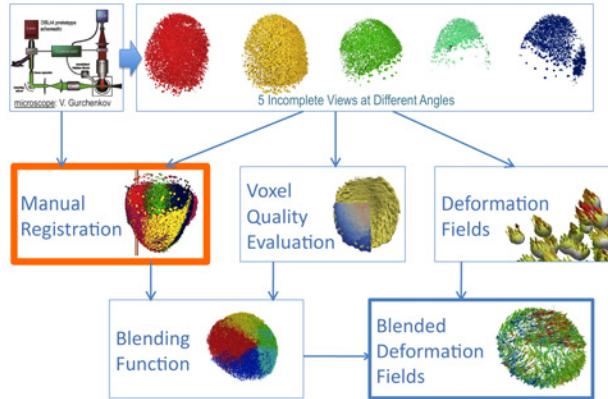


Figure 7.9: The manual registration component. This process aims at inferring the parameters required for a spatial registration of the five angles of view.

section 7.2, we use the following conventions to name the axes of each view: the axis from the camera to the embryo is the z-axis or “depth”, the horizontal axis of the light plane is the x-axis, and the vertical axis of the light plane is the y-axis.

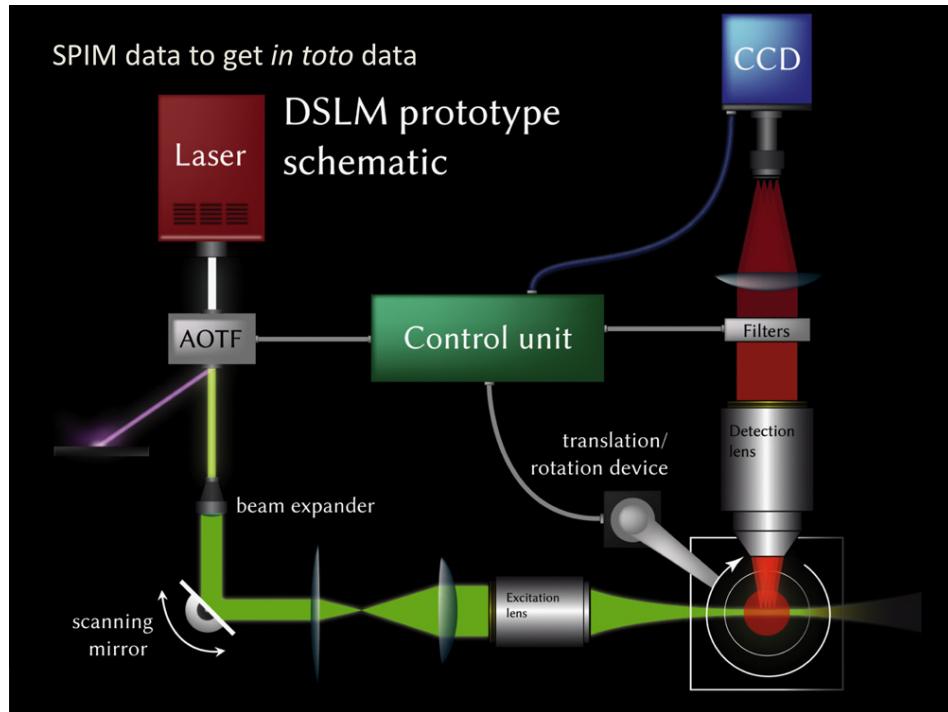


Figure 7.10: Digital Scanning Light Microscope (DSLM) setup. Diagram after P.J. Keller, E. Stelzer et al. [277], adapted by V. Gurchenkov.

Reconstructing an *in toto* view of the embryo requires the spatial fusion of the five views within a reference frame. While sophisticated methods of non-rigid spatial registration based on raw image processing and “wavelets” are currently being developed by our team [307], we decided to design our own for faster use. Here, we rely on the knowledge of the rotating motion of the microscope plate to geometrically infer a rigid spatial registration. Moreover, we assume that the registration parameters remain constant throughout the imaging process since the embryo is not moving in the agarose chamber and the rotating plate executes precise and controlled displacements. Thus we use a single time step to calculate the parameters, which will be generalized to all other time steps.

The embryo is mounted onto a cylindrical agarose tube. The tube is fixed on a table that can rotate around the cylinder’s central axis and translate it. As the embryo never passes exactly through that central axis, the software controlling the motion of the table must be calibrated before the time-lapse imaging starts. The calibration phase amounts to the following process (Fig. 7.11): the operator must center the embryo in the light plane. Once the position is validated, the software performs a 90 degree rotation and the operator must center again the embryo in the light plane. Based on the translation vector created by the operator’s action, the software computes the relative position of the embryo with respect to the cylinder’s center. However, the computed relative position is only an approximation of the real relative position because the operator’s estimation is never exact. This is due to the fact that s/he only sees a projection along the z-axis, and as the embryo rotates, the aspect on the screen changes.

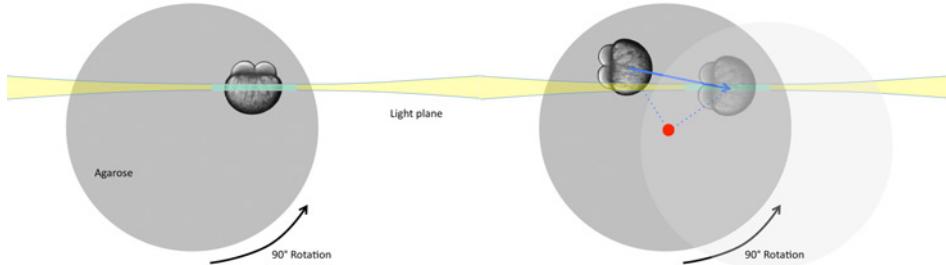


Figure 7.11: Schema of the calibration phase. Left: The operator centers the embryo in the light plane. Right: To decipher the relative position of the embryo with respect to the center of rotation, the software performs a 90 degree rotation, then the operator must translate the embryo back into the light plane (blue vector).

We have programmed a graphical user interface to manually correct the estimation error and determine the optimal relative position (Fig. 7.12). The two parameters controlled by the user are the coordinates of the correction vector. As

the user moves this vector, s/he can directly observe the motion of the five views computed in real time, hence determine the optimal correction vector. For this task, the nuclei of the margin (yolk syncytial layer), which are separated from the mass of the deep cells, offer useful landmarks for a precise determination of these vector coordinates, i.e. the best match (Fig. 7.13).

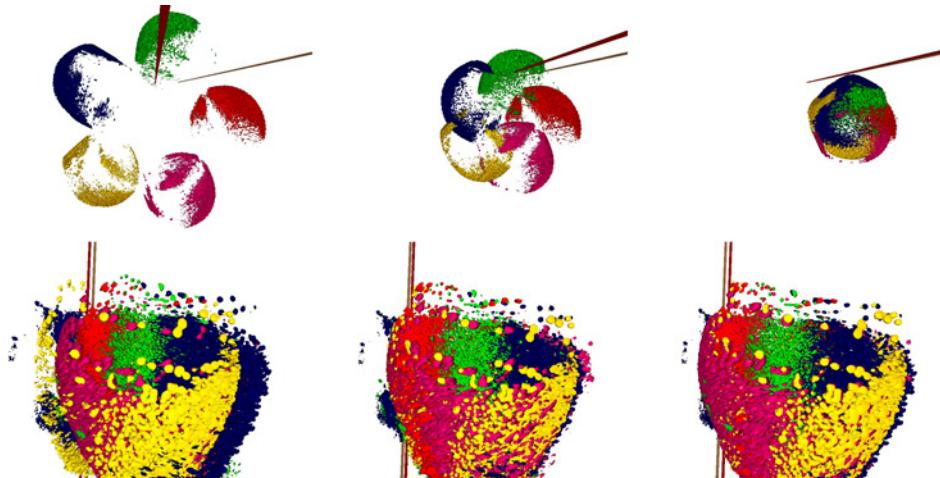


Figure 7.12: Graphical user interface designed to infer the estimation of the center of rotation. The red line passes through the center of rotation determined by the calibration phase. The brown line passes through the corrected center of rotation. As the user moves the brown line, the five angles of views are repositioned in real time, allowing to infer the best estimation of the center of rotation.

After the five views have been registered for one time step, the same registration parameters will be applied to the whole time-lapse imaging data in the other components of the workflow (Fig. 7.7). However, as it can be noticed in Fig. 7.14, the optical deformation induced by the perturbation of the light path through the embryonic tissue causes the result to be rather suboptimal. Other components of the workflow (“Voxel Quality Evaluation” and “Blending Function” below) will take care of removing these defects in each view.

7.2.2 Voxel Quality Evaluation

The voxel quality component runs independently from the manual registration component. Its goal is to determine what spatial region of the volume shows the best signal quality, in each view and at every time step. The actual output of this component is to attribute a quality score to each voxel of the volume. The rationale here is that the farther the light travels through the embryo, the more degraded the signal becomes. The light beam through the embryo is split into two orthogonal

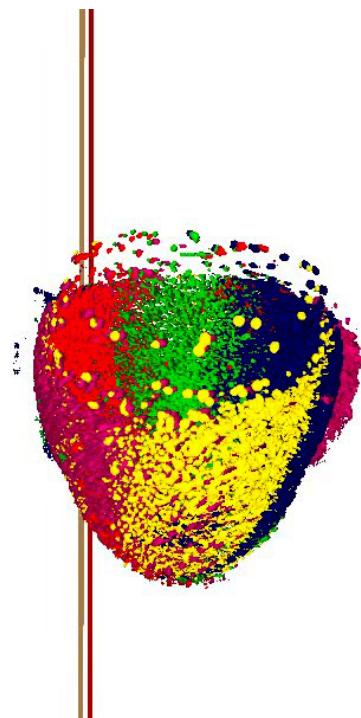


Figure 7.13: Best match for the rigid manual registration.

The marginal yolk nuclei visible on top of the embryo are used to estimate the correct registration parameters.

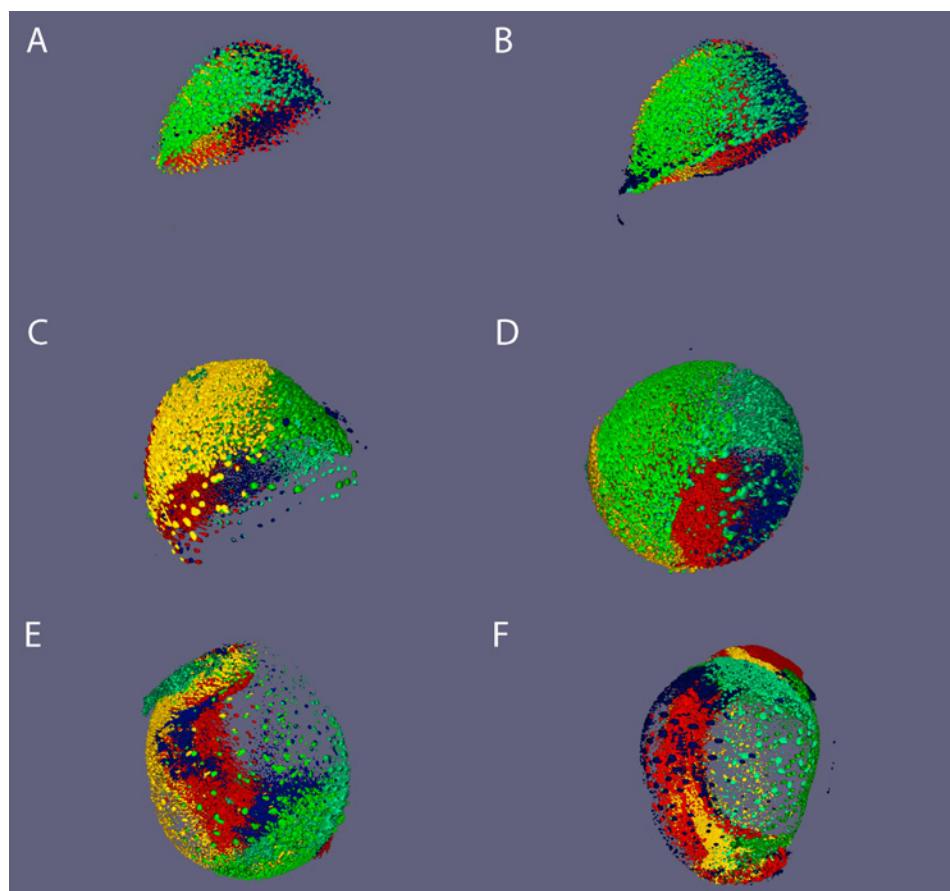


Figure 7.14: The five DSLM views superimposed by rigid registration. These views are identical to the ones shown in Fig. 7.8. A: Sphere stage. B: 40 percent epiboly. C: 75 percent epiboly. D: 90 percent epiboly. E: 8-somite stage. F: 12-somite stage.

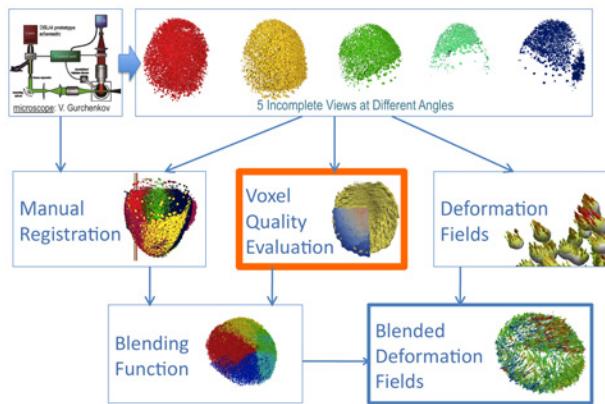


Figure 7.15: The voxel quality component. This process aims at quantitatively evaluating the local image quality for each angle of view.

beams: the continuation of the original light beam from the laser, and a new light beam emitted by the fluorescent molecules of the embryo toward the camera. For the voxel quality component, the relevant part of the light beam is inside the embryo. We performed a geometrical computation to determine the length of the light beam from each voxel of the volume to the physical border of the embryo. For the evaluation, we first calculated the border and then the beam length. This task can be summarized as follows:

- 1. To eliminate the low amplitude background noise, the average voxel intensity of the volume was computed to determine the isosurface of the nuclei's shapes (Fig. 7.16A). This isosurface is a mesh of triangles.
- 2. After a random decimation of the isosurface triangles to decrease the number of vertices, small spheres were centered on these vertices and resized in order to produce a contiguous embryo surface without holes. Then, a binary mask of the image was extracted to store the outer shape of the embryo.
- 3. We eliminated spurious signal traces (blue arrows in Fig. 7.16B) that may have subsisted outside of the embryo shape by keeping only the closed region that had the largest volume.
- 4. The spheres creating an embryo larger than desired, we eroded the volume by an amount equivalent to the radius of the spheres to obtain the actual physical border of the embryo. The reduced volume was stored as a binary data array with value 1 (resp. 0) for voxels inside (resp. outside) the border (Fig. 7.16C).
- 5. Since each view was oriented with its z-axis along the camera path and x-

axis along the illumination plane, it was now straightforward to compute both distances from the border by counting the number of 1's along those axes (Fig. 7.16D,E). The voxels that were outside of the border received an eliminatory score as they could not be used for the registration task (Fig. 7.16F).

- 6. Finally, we also eliminated the voxels located beyond a cutoff distance (about $200\mu\text{m}$) in both directions as they belonged to parts of the image that were too degraded to be used (Fig. 7.16G,H,I).

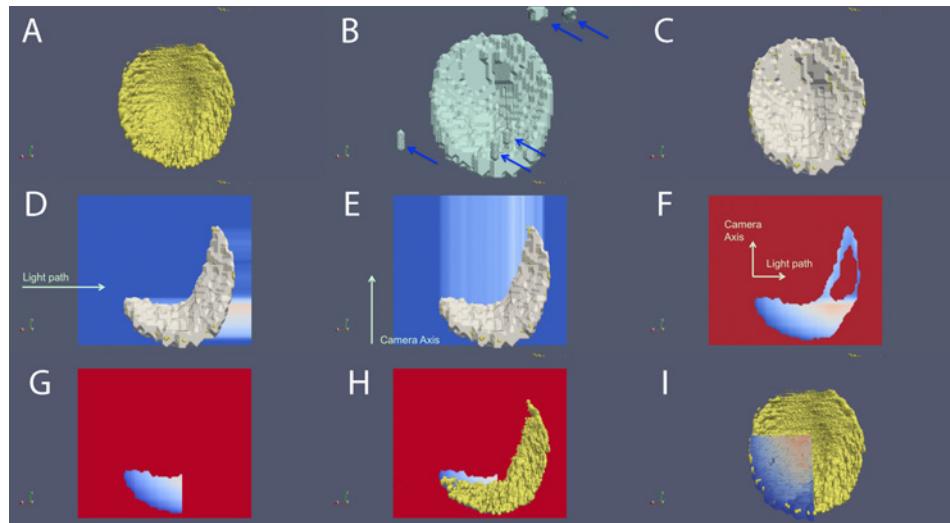


Figure 7.16: The various steps of the voxel quality evaluation component. See text.

7.2.3 Blending Function

The objective of this task is to use both the manual registration component and the voxel quality evaluation component described in the previous subsections to build a blending function. For each registered voxel, this function selects among all five views which one will be used in the *in toto* reconstruction. In other words, the blending function produces a 3D integer volume in which each voxel contains the identifier (id, from 1 to 5) of the most relevant angle of view according to its quality evaluation. Knowing the coordinates of the center of rotation, we were able to precisely position all five views within a common reference space (Fig. 7.18, left). In this space, we visited each voxel and checked for each projected angle of view whether it was located inside the embryo border evaluated in the previous component (Fig. 7.18, right). Then, we attributed to this voxel the id of the angle of view under which it had the best quality, or 0 if it was lying outside of all borders. This produced the integer mask shown in Fig. 7.19. Note that instead of the

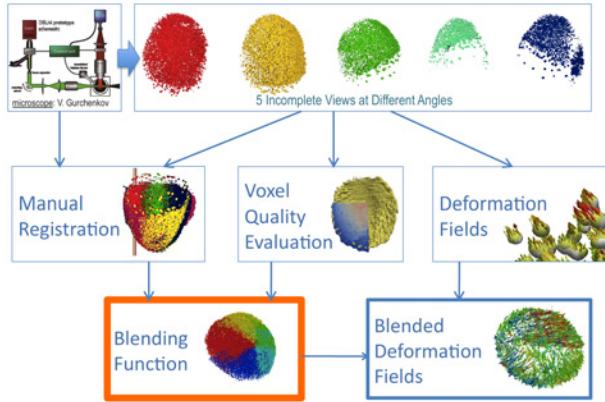


Figure 7.17: The blending function component. This component builds a function determining what original angle of view has the best quality score for each voxel in the registered space.

maximum voxel quality, we could also have used a weighted average of all angles. An output of the blending function applied to the raw voxel intensities is shown in Fig. 7.20, providing a first example of *in toto* registration of the embryo. In our workflow, however, we applied the blending function to the “deformation fields” calculated in the next component.

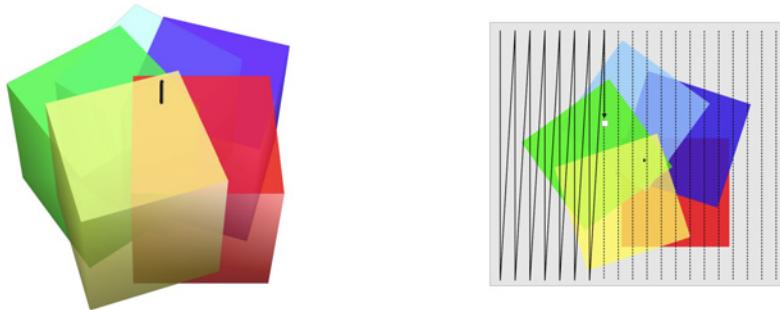


Figure 7.18: Relative positions of the five views in the registered space. Left: The five cubes represent the original reference frames of the five different views. Right: A 2D illustration of the sequential visit of the registered voxels.

7.2.4 Deformation Fields

The goal of the deformation field component is to reconstruct the spatial deformation that relates two 3D volumes produced by the DSLM at consecutive time steps under the same angle of view (one of the five; Fig. 7.22), which we denote here by V and

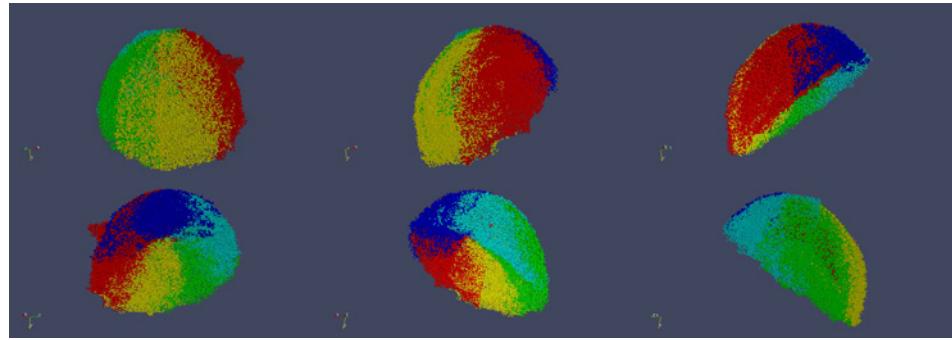


Figure 7.19: Five-valued blending function mask. In the registered space, the color of each voxel represents the id of the original angle of view under which it has the best quality.

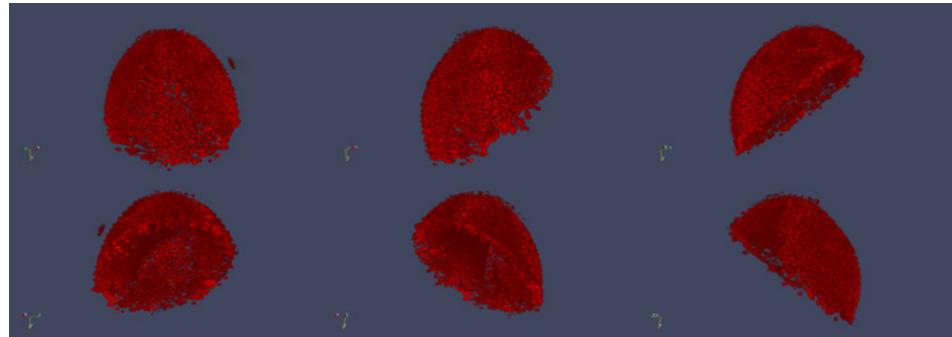


Figure 7.20: Output of the blending function applied to the raw voxel intensities.

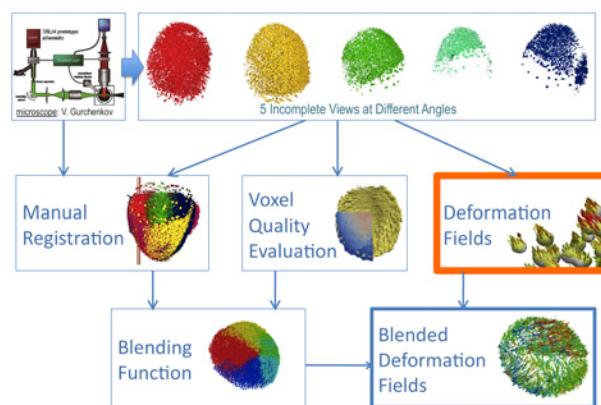


Figure 7.21: The deformation field component. This component produces a 3D vector field that indicates the direction and amplitude of local displacements in each voxel.

V' . The notion of image deformation is close to the notion of “optical flow”, as its goal is to compute the image motion between consecutive images. We employed a non-rigid registration method called “Demons Registration”: it acts as if voxel-sized “demons” were pulling and pushing the voxel of volume V' along the local gradient of voxel intensity in volume V . This method is iterative and the demons run for a predefined number of steps, or until some matching criterion is reached. The implementation that we used here is based on Thirion’s algorithm [308]. Fig. 7.23 shows the superposition of the isosurfaces of V (in white) and V' (in yellow) at the marginal region of the zebrafish embryo. As the deformation field was computed in the entire images of V and V' , and not only at the position of the nuclei, we had to filter out spurious vectors from the resulting field. The vectors belonging to voxels located inside the nucleus envelopes (corresponding to another isosurface obtained by thresholding the voxels at twice the mean volume intensity) were retained, while the others were discarded. We calculated the deformation fields for each angle of view separately. The next component below will take care of blending them together.

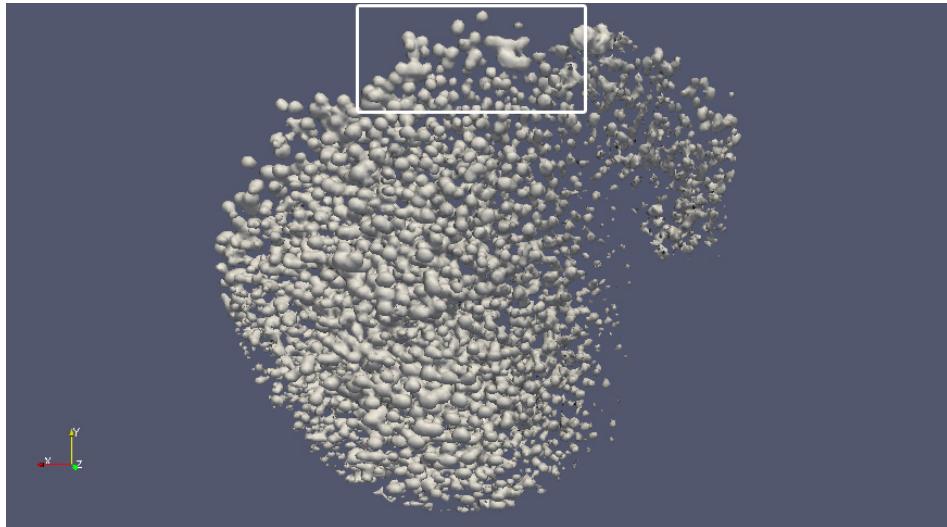


Figure 7.22: One volume V under one of the five angles of view. The white grains represent the envelopes of the nuclei calculated by thresholding the voxels at twice the mean intensity of V .

7.2.5 Blended Deformation Fields

Finally, the blending function obtained in Section 7.2.3 is applied to the five deformation vector fields obtained in Section 7.2.4. Each voxel of the reference volume simply receives the deformation vector from the angle of view that has the best

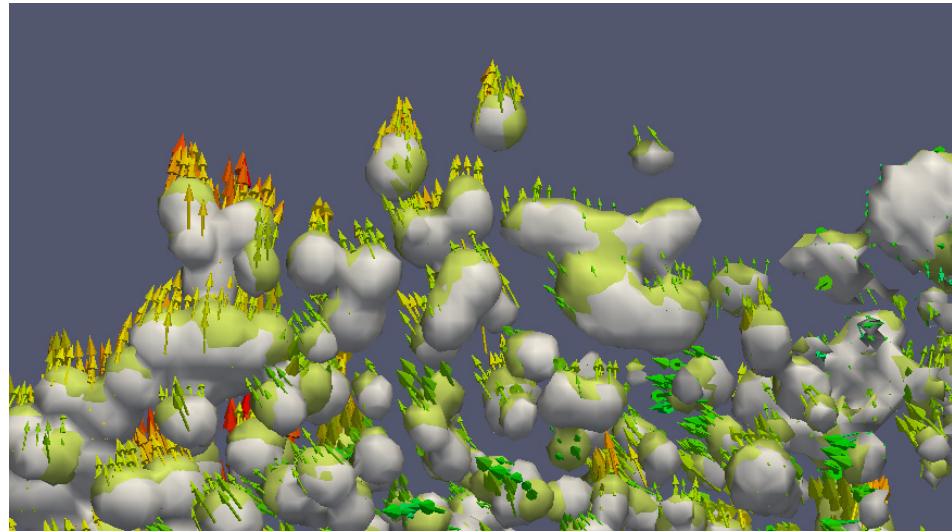


Figure 7.23: Deformation field between two volumes V and V' at consecutive time steps. The white isosurface represents V , the yellow isosurface represents V' . Not all deformation arrows are displayed.

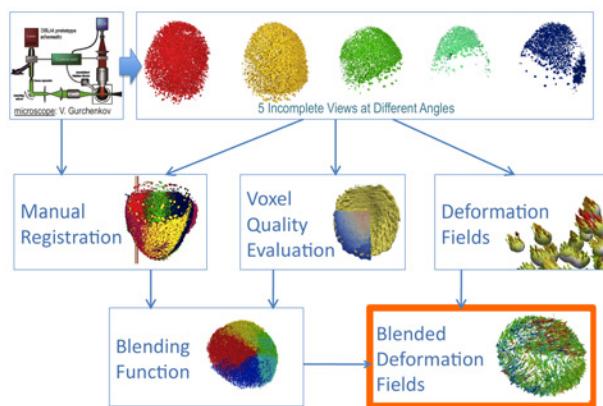


Figure 7.24: The blended deformation field component. This component combines the blending function with the 3D deformation vector field.

quality evaluation score. This produces our final output, the “blended deformation field”.

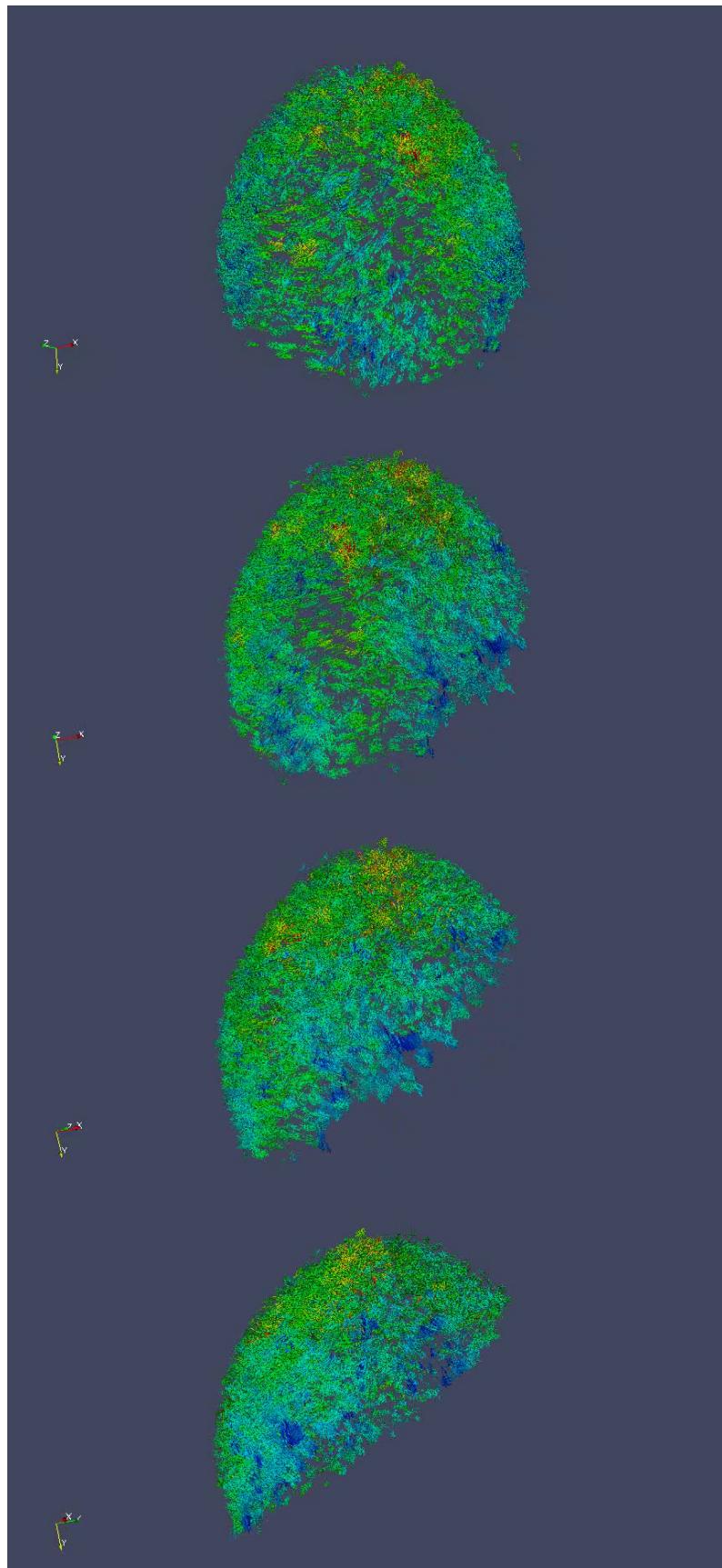


Figure 7.25: Blended deformation field at a given time step viewed from four different angles. (Note that these angles do not correspond to the original angles of view.)

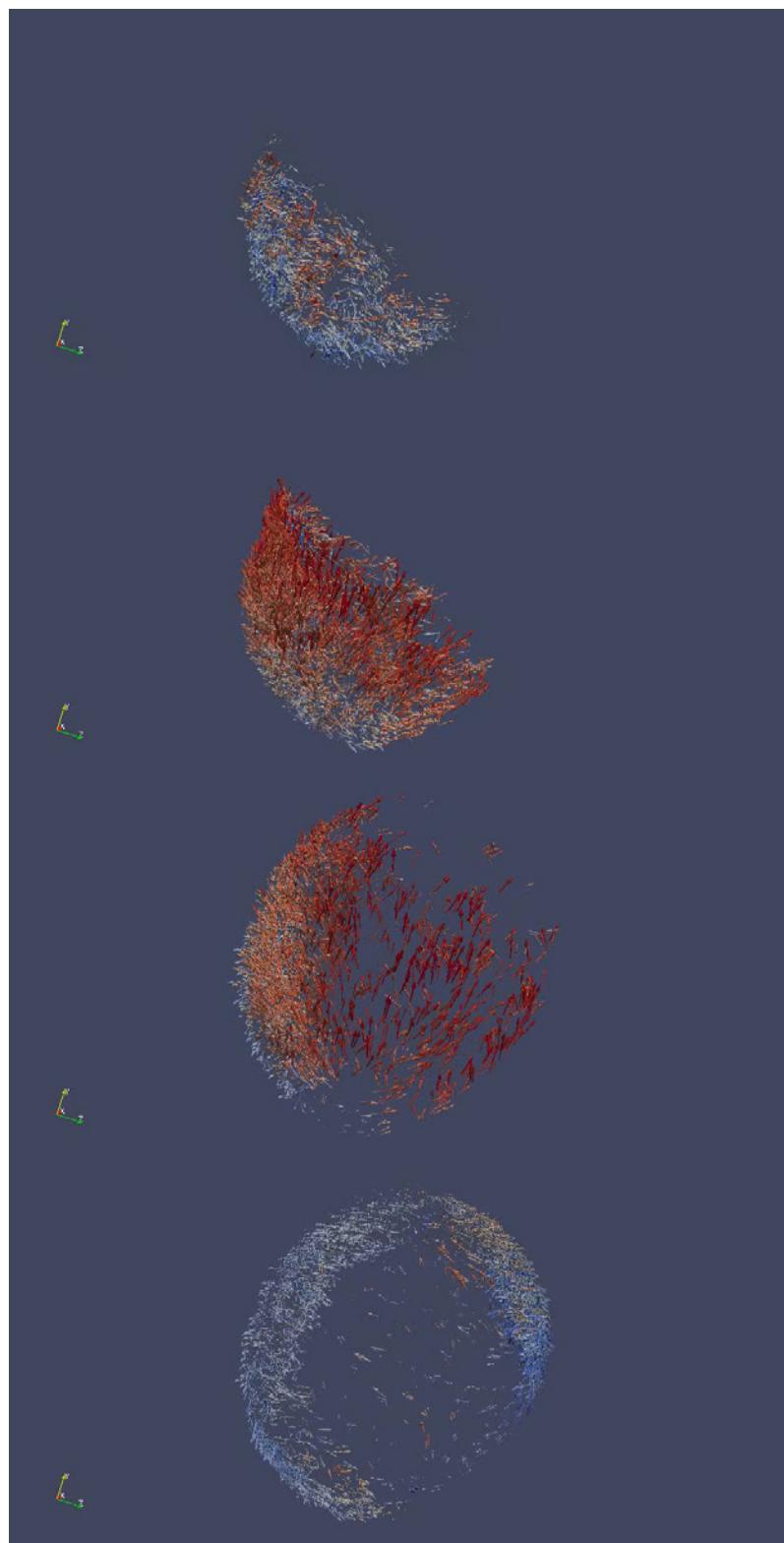


Figure 7.26: Blended deformation field at successive time steps viewed from the same angle.

CHAPTER 8

Simulated Reconstruction and Model Validation Through Case Studies

The zebrafish early development is the site of multiple morphogenetic events that illustrate the links between microscopic cell behaviors and macroscopic deformations. We chose to examine more closely and treat six of these events by modeling and simulation. Our work is reported here in the form of independent exercises, or *case studies*. The idea is that the general mechanogenetic model described in Chapters 3-5 can be customized and applied to various episodes of embryogenesis, in particular to the zebrafish for which we have imaging data as validation.

Consequently, this chapter is organized in six sections:

- **Case Study 1: Investigating the Yolk Biomechanical Properties** (Section 8.1). We explore the viscoelastic properties of the yolk cell at the zygote stage. These properties are assessed by a mechanical experiment where a spherical bead is pushed into the yolk's volume, and the subsequent relaxation of the yolk's surface is measured.
- **Case Study 2: Cell Proliferation Rate Along the Cell Lineage** (Section 8.2). The pattern of cell cycles of the deep cells, which start proliferating between the yolk and the enveloping layer, is investigated during the first ten hours of zebrafish development. The question of the regularity of these cycles and the law they follow (e.g. geometrical or arithmetic) is examined for cell cycles 10 to 14.
- **Case Study 3: Shaping the Zebrafish Blastula** (Section 8.3). The shape of the blastula is a consequence of the mechanical interactions among blastoderm cells and between the blastoderm and the yolk. During the developmental period, the enveloping layer (EVL) compartment is formed. We present the specificities of the EVL structure and its interactions with the deep cells and the yolk. Our strategy to evaluate these interactions are also presented. This section reviews ongoing research that is not fully reported here.
- **Case Study 4: Cell Behaviors in the Enveloping Cell Layer Compartment** (Section 8.4). The EVL is hypothesized to behave as a passive

tissue reorganizing its cell neighborhood under an external tension exerted at the EVL-yolk margin. Specific geometrical and topological measures are introduced and used to compare a model of epithelial growth and proliferation with live specimen observations during epiboly.

- **Case Study 5: Intercalation Patterns** (Section 8.5). Cells intercalate during epiboly, causing the overall embryo shape to flatten. We investigate various parameters involved in this process, from an active protrusion behavior displayed by the deep cells to the mechanical resistance of the EVL and the EVL-yolk margin.
- **Case Study 6: Gastrulation** (Section 8.6). We present how we envision the different morphogenetic events occurring during gastrulation: internalization, convergence-extension, and late epiboly. We also present measures adapted to their characterization. This section reviews ongoing research that is not fully reported here.

A note on the choice of parameters

Each case study follows the same methodological and experimental setup. First, the phenomenon and biological question are introduced. Next, we describe the specific ingredients of the model (cellular types and mechanisms at play). Then, the relevant parameters are presented and their effect on the adequacy of the model with respect to the observations is explored according to the methodological workflow introduced in Section 1.1.3 (Fig. 8.1). Finally, results are critically assessed and discussed.

Naturally, since the computational time required to sweep the “entire” parameter space (or a good portion of it) would be much too large, because of the high number of dimensions and possible values, we conduct parameter exploration in two phases:

- 1. An “intuitive” phase: At first, we explore parameter space manually with a trial-and-error approach, where diverse sets of parameters are tested and evaluated by visualizing the simulated phenotype. The efficacy of this phase depends on the experience of the modeler, her/his understanding of the model, and the objectives of the case study. Practically, a graphical user interface (GUI) supporting the simulations constitutes an indispensable tool to rapidly select the most relevant parameters. It makes it possible to immediately observe the consequence of their variation on the developmental trajectories.
- 2. A “systematic” phase: Once a set of candidate parameters has been identified, the process of simulating and evaluating the fitness of each phenotype is automated. On each parametric dimension, extremum values of the interval of variation are set, as well as the number of sample points within that interval. Then, an automated evaluation loop produces a fitness landscape or

a “phase map” associating each point in the parameter space to a measure of the embryo. These fitness landscapes are analyzed and discussed.

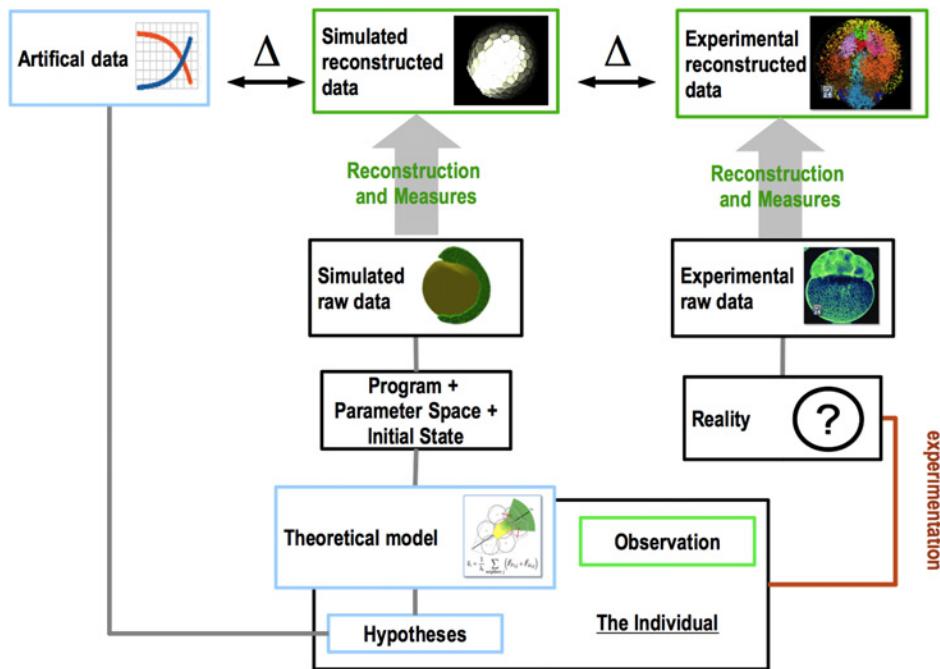


Figure 8.1: Chapter 8 involves all the components of the methodological workflow introduced in Section 1.1.3.

8.1 Investigating the Yolk Biomechanical Properties

We assume that the early development of the embryo depends on *viscoelastic properties* of the yolk. Using our simulation framework, our aim is to identify these properties by assessing their effect on later morphogenesis. In particular, we expect that yolk properties help maintain the integrity of the whole embryo, for example that the yolk should resist pressure exerted by the cells. At the same time, it should also undergo successive transformations by flattening at the sphere stage, and protruding at the doming stage. In addition, its active role in gastrulation is highly debated [103]. Several of these aspects will be investigated in the following sections. We show here that minimal viscoelastic properties are sufficient to account for some of the embryo’s phenotypic features throughout cleavage and gastrulation.

8.1.1 Hypotheses and Model

Our model is based on state-of-the-art observations. The zebrafish zygote is a large cell filled with cytoplasm and lipid droplets. These two fluids start to segregate before the first divisions, then the cytoplasm is gradually sucked into the dividing cells (Fig. 8.2). A lipid bilayer membrane progressively separates the cells at the animal pole from the lipid droplets, forming an independent structure called the *yolk cell*. After the 9th division cycle, the marginal deep cells divide with one daughter cell fusing with the yolk, forming a structure just beneath the cells called the *yolk syncytial layer* (YSL). Another layer, the *yolk cortical layer* (YCL), completely envelops the yolk (See below).

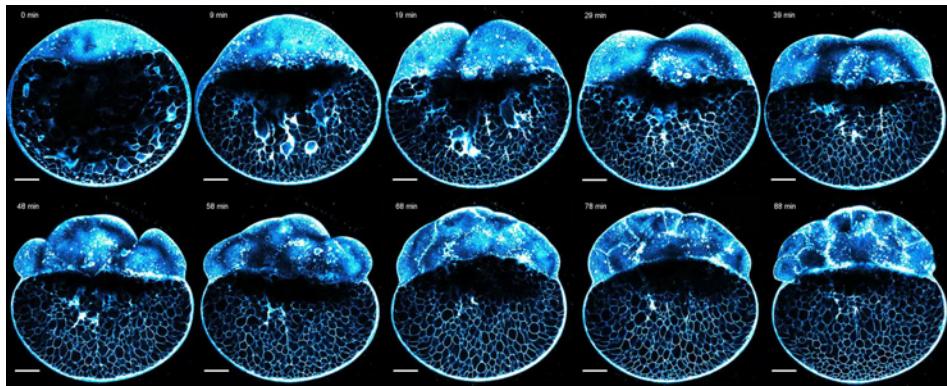


Figure 8.2: Third harmonic generation (THG) imaging of the zebrafish embryo adapted from [254]. Snapshots from the movie S.1 going from the one-cell stage to the 64-cell stage. Developmental timing (min) indicated top left. Scale bar (100 microns) indicated bottom left.

In this context, from the zygote stage, we make the approximation that the cells and the yolk are independent structures—but to capture their dynamical mechanical interactions, we model both of them in the same particle-based paradigm. Furthermore, we hypothesize that it is relevant to distinguish between an external “yolk membrane” (ym), which is composed of the YSL and YCL together, and the inner lipid droplets, or “yolk interior” (yi). Thus we use three types of particles: one for the cells (displayed in green in Fig. 8.3) and two for the yolk: the membrane particles (in yellow) forming a 2D viscoelastic layer, and the interior particles (in red) representing the lipid droplets confined by the external layer. We expect the external membrane to keep its surface topology, whereas the internal structure can rearrange itself. This is a qualitative interpretation of imaging that provides plausible properties (S.1).

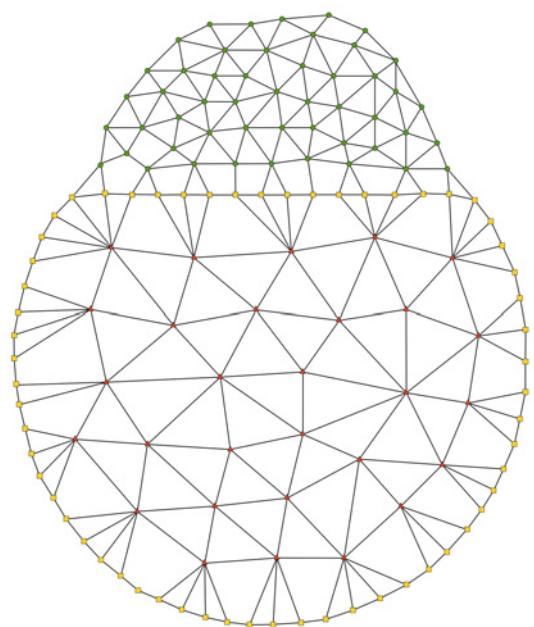


Figure 8.3: Abstract schema of the particle types in the modeled embryo. Green particles represent the cells, yellow particles the external “yolk membrane” (ym), and red particles the inner lipid droplets or “yolk interior” (yi).

The yolk membrane (ym)

In Fig. 8.3, the yellow particles materialize the yolk membrane. Note that, whether we are talking of the YSL or the YCL, this “yolk membrane” is not only composed of a lipid bilayer but also a cytoskeleton that underlies it (acto-myosin network, micro-tubules, etc.). Here, the membrane topology is defined by a geodesic dome derived from an icosahedron, one of the five “Platonic solids”, which remains invariant through time. An icosahedron’s surface is composed of 20 identical equilateral triangles, 12 vertices and 30 edges. This topology is refined by subdividing each triangle into four new identical equilateral triangles. The new vertices are projected on the sphere. This process is repeated 4 times to approximate a spherical shape with 5120 triangles and 2562 vertices (Fig. 8.5). Although other Platonic solids could have been used without much difference, the icosahedron offers a better trade-off between stability (its elementary faces are triangles) and volume maximization (only the dodecahedron occupies a slightly larger volume, but its faces are pentagonal). In our geodesic dome, each vertex except the original 12 has 6 direct neighbors, the original vertices having only 5 direct neighbors. These direct neighbors are labeled “rank-1” neighbors. For each ym particle i , this neighborhood is denoted by $\mathcal{N}_i^{\text{ym},1}$.

The forces exerted between ym particles are pure elastic forces. Each rank-1 neighborhood relationship is materialized by a linear spring, so that between two ym particles (i,j) , the force is simply:

$$\vec{F}_{ij}^{\text{ym}} = -k_{\text{ym}}(r_{ij} - r_{ij}^{\text{eq}})\vec{u}_{ij}$$

where k_{ym} is the stiffness constant controlling the ability to deform the neighborhood link under a given force applied to it (a higher k_{ym} creating a stiffer resistance), and r_{ij}^{eq} is the resting length of the neighborhood link (i.e. its length if no force is applied). Each ym-to-ym link has its own characteristic resting length. It is defined by the length of the link when the ideal geodesic dome is projected on a sphere whose volume is equal to the desired total yolk volume V_Y , i.e. the radius is $R_Y = (3V_Y/4\pi)^{1/3}$. Concretely, the dozen 5-neighbor vertices have shorter resting lengths than all the other 6-neighbor vertices. Without this rescaling, the equilibrium shape would converge toward the original Platonic solid and not toward the sphere.

Since the ym neighborhood topology is invariant (ym particles do not “swarm”, so the geodesic links remain the same), there is no plasticity in the yolk membrane alone. The original shape is recovered after local mechanical deformation (local pushing or pulling). Beyond a certain level of stress, however, the membrane mesh may get entangled and erratic behavior may occur. This is illustrated by an “abnormal” lip formed by the yolk membrane when ym particles are only interacting with their rank-1 neighbors (Fig. 8.4A). We hypothesize that increasing the ym stiffness and the spatial range of particle interactions would prevent this undesirable feature. When constraining ym particles to rank-1 and rank-2 interactions, the blastoderm-

yolk margin remains smooth (Fig. 8.4B). Rank-2 neighbors of a yolk membrane particle are defined as the set of rank-1 neighbors plus their rank-1 neighbors (Fig. 8.5F). All the corresponding links are materialized by linear springs with the same stiffness coefficient k_{ym} and resting lengths r^{eq} defined by the same procedure as rank-1 neighbors. The rank-2 neighborhood set is denoted by $\mathcal{N}_i^{ym,2}$.

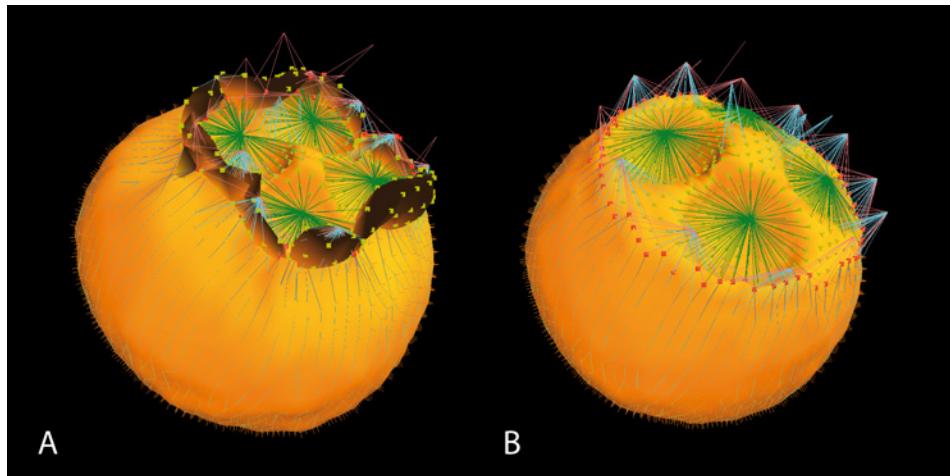


Figure 8.4: Yolk membrane behavior depending on its particles' topological interactions. A: rank-1 interactions only. B: rank-1 + rank-2 interactions.

So far, we have idealized the yolk membrane as an “empty balloon”. It must now be filled with a viscous fluid. In the real yolk, if the yolk membrane is damaged, inner material is expelled out of the membrane. This indicates that the inside pressure is higher than the outside. We hypothesized that this difference of pressure is due to the cortical tension exerted by the yolk membrane and introduce an additional parameter: c_{r1} , which is a dimensionless coefficient reducing the resting lengths within the yolk membrane. The above force equation then becomes:

$$\vec{F}_{ij}^{ym} = -k_{ym}(r_{ij} - c_r r_{ij}^{eq}) \cdot \vec{u}_{ij}$$

The yolk interior (yi)

The yolk membrane as modeled above has the ability to absorb and recover from tangential load but can hardly resist any load applied in a non-tangential manner. Inspired by the lipid droplets contours revealed by THG signals in Fig. 8.2, we fill the yolk interior (yi) with N_{yi} identical particles of volume V_{yi} and radius R_{yi} (red points in Fig. 8.3). Practically, $N_{yi} = 500$ particles were added in our simulation, which is less than the number of lipid droplets estimated from the 3D data, and represents a coarse-grained approximation. We assume that the ym particles’ radius is equal to the yi particles’ radius, so that the volume occupied by the N_{yi} particles

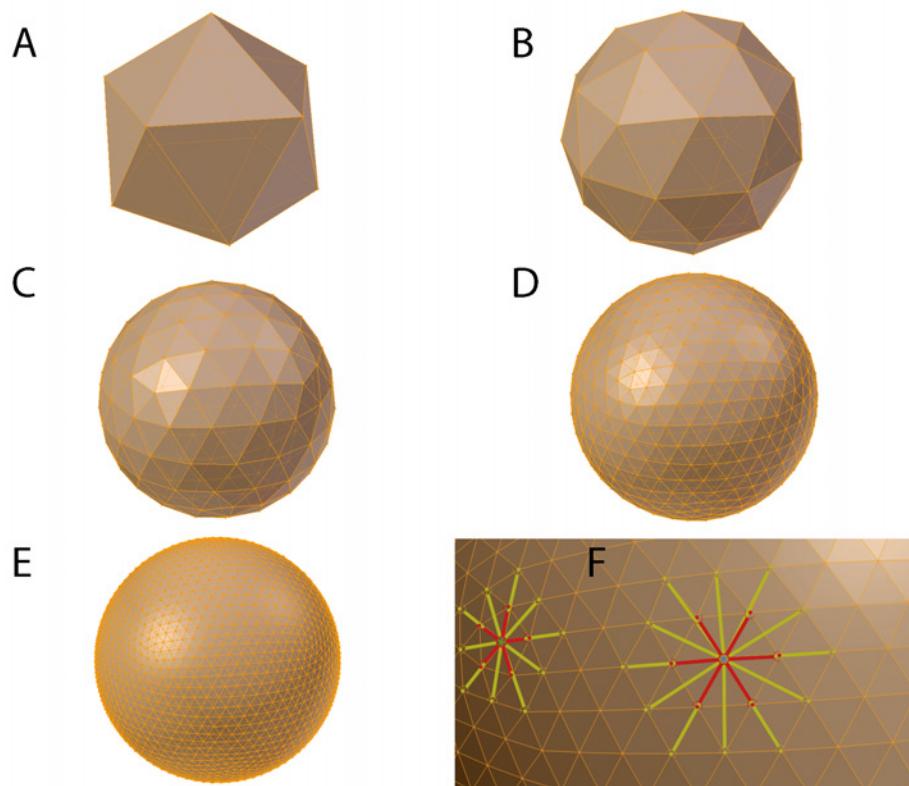


Figure 8.5: The structural topology of the yolk membrane is a geodesic dome. A: Icosahedron. B,C,D,E: iterative subdivision of the original icosahedron to obtain the geodesic dome. F: the right vertex is a regular geodesic vertex and the left vertex is one of the 12 original icosahedral vertices. Rank-1 neighbors link in red, rank-2 neighbors link in yellow.

is equivalent to the volume of a sphere minus the thickness of the ym layer, i.e. whose radius is $R_Y - R_{\text{ym}} = R_Y - R_{\text{yi}}$.

$$N_{\text{yi}} \frac{4}{3} \pi R_{\text{yi}}^3 = \frac{4}{3} \pi (R_Y - R_{\text{yi}})^3 \Leftrightarrow R_{\text{yi}} = \frac{R_Y}{1 + \sqrt[3]{N_{\text{yi}}}}$$

and with $R_Y \gg R_{\text{yi}}$, we get $R_{\text{yi}} \approx R_Y (N_{\text{yi}})^{1/3}$. We apply the same rules as the ones described for cell-cell interactions in Chapter 3 to ym-yi and yi-yi pairs of particles. As in Section 3.2.2, the neighborhood of interactions is recalculated at every time step of the simulation, first through a metric criterion then by topological selection. This adaptive topology provides to the yolk a certain plasticity through passive rearrangement of the yi particles. Both ym-yi and yi-yi interaction forces have the same expression as the passive attraction/repulsion forces between deep cells particles (see Section 3.2.3):

$$\vec{F}_{ij}^y = A_{ij} \cdot \vec{F}_{ij}^{y,\text{lin}} = \begin{cases} -w_{\text{rep}}^y (r_{ij} - r_{ij}^{\text{eq}}) A_{ij} \cdot \vec{u}_{ij} & \text{if } r_{ij} < r_{ij}^{\text{eq}} \\ -w_{\text{adh}}^y (r_{ij} - r_{ij}^{\text{eq}}) A_{ij} \cdot \vec{u}_{ij} & \text{if } r_{ij} \geq r_{ij}^{\text{eq}} \text{ and } r_{ij} < r_{ij}^{\text{max}} \\ \vec{0} & \text{if } r_{ij} \geq r_{ij}^{\text{max}} \end{cases}$$

where “y” stands for either ym-yi or yi-yi types of interactions, and w_{adh}^y and w_{rep}^y are unique yolk-specific stiffness coefficients common to both types.

Parameter space for the biomechanical properties of the yolk

In summary, putting together the ym and yi force expressions above, the equation of motion of the N_{ym} particles of the yolk membrane reads:

$$\vec{v}_i^{\text{ym}} = \frac{1}{\lambda_0 R_{\text{ym}}^2} \left(\sum_{j \in \mathcal{N}_i^{\text{ym},2}} -k_{\text{ym}} (r_{ij} - c_r r_{ij}^{\text{eq}}) \cdot \vec{u}_{ij} + \sum_{j \in \mathcal{N}_i^t} \vec{F}_{ij}^y \right)$$

and the equation of motion for the N_{yi} particles of the yolk interior is:

$$\vec{v}_i = \frac{1}{\lambda_0 R_{\text{yi}}^2} \sum_{j \in \mathcal{N}_i^t} (\vec{F}_{ij}^y)$$

where neighborhood $\mathcal{N}_i^{\text{ym},2}$ contains only ym-type particles and \mathcal{N}_i^t contains both ym and yi types. This system of equations is controlled by the following parameters: λ_0 , w_{adh}^y , w_{rep}^y , k_{ym} , c_r , the simulation time step Δt_s and a number of other background parameters such as R_{ym} , R_{yi} , a , and c^{max} (in the expression of A_{ij}), which were a priori fixed or pre-calibrated when designing the simulation. The simulation, parameter space exploration and validation process described in the next subsection will allow us to estimate realistic values for the following ratios:

- $\bar{w}_{\text{adh}}^y = w_{\text{adh}}^y / \lambda_0$

- $\bar{w}_{\text{rep}}^{\text{y}} = w_{\text{rep}}^{\text{y}}/\lambda_0$
- $\bar{k}_{\text{ym}} = k_{\text{ym}}/\lambda_0$

the last two undetermined or “free” parameters being the yolk membrane surface scaling parameter c_r and the simulation time step Δt_s .

8.1.2 Simulation, Parameter Space and Validation

We expect the yolk to exhibit a specific behavior of *deformation* under external stress followed by a behavior of *resilience* through recovery. In order to assess the relevance of these properties, we designed an experimental protocol applied to both live specimens and simulated ones. As in most of the case studies explained in this Chapter 8, confronting measurements in both cases is our strategy to explore the parameter space, with the goals of:

- qualitatively discussing the fitness landscape, and
- finding optimal parameters.

Experimental protocol: live specimen

A spherical 1-cell-stage embryo lying on its side in an agarose mold, a glass bead is pressed on the lateral side of the embryo for a period of time ΔT , starting at t_0 . The force is applied along an axis parallel to the agarose surface and, on the opposite side, the embryo is blocked by an agarose wall. Pressure is applied to the bead until it enters into the yolk by approximatively 40

Qualitatively, the yolk is deformable with the yolk membrane surrounding the bead in a crater-like structure somewhat larger than the bead diameter. After removing the bead, the yolk quickly recovers its shape. The following measurements allow us to quantify this process. We measured every time interval $\Delta t_y = 0.8$ second the normalized length of the yolk along the compression axis, defined as the ratio of the deformed embryo’s diameter L over the relaxed embryo’s diameter L_{rel} before compression.

Experimental protocol: simulation

The simulated experimental protocol mimics the live one described above. The algorithmic implementation of the model integrates an artificial spherical bead which is forced to move toward the center of the yolk. Yolk particles, whose original positions are now occupied by the bead, are displaced by being projected on the bead’s surface. Similarly to the experimental live protocol, we specify the force necessary to induce the particles’ motion but impose their displacement. The ball is removed when it has sunk to a depth equal to 40

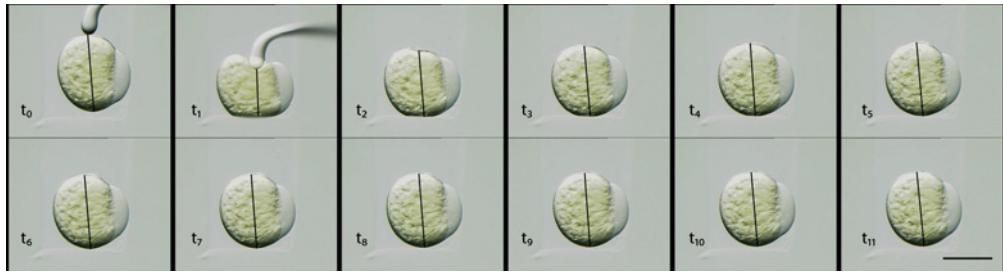


Figure 8.6: Experimental deformation of the egg yolk. 1-cell stage, lateral view, animal pole to the right. t_0 : Onset of the deformation. t_1 : Maximum deformation. Measurement is realized every 0.8 second. We measure the distance L between the area pointed by the tip of the bead and the opposite side (black line). Scalebar 500 microns.

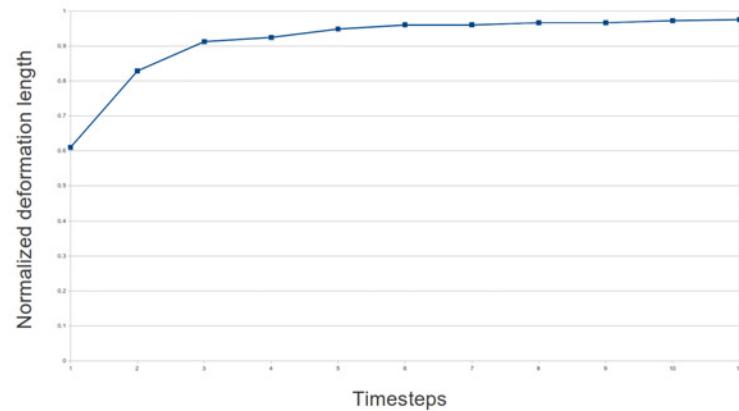


Figure 8.7: Plot of the normalized deformation length as a function of time. The shape recovery is asymptotic toward the normalized length at t_0 . Full recovery is nearly achieved by 8 seconds.

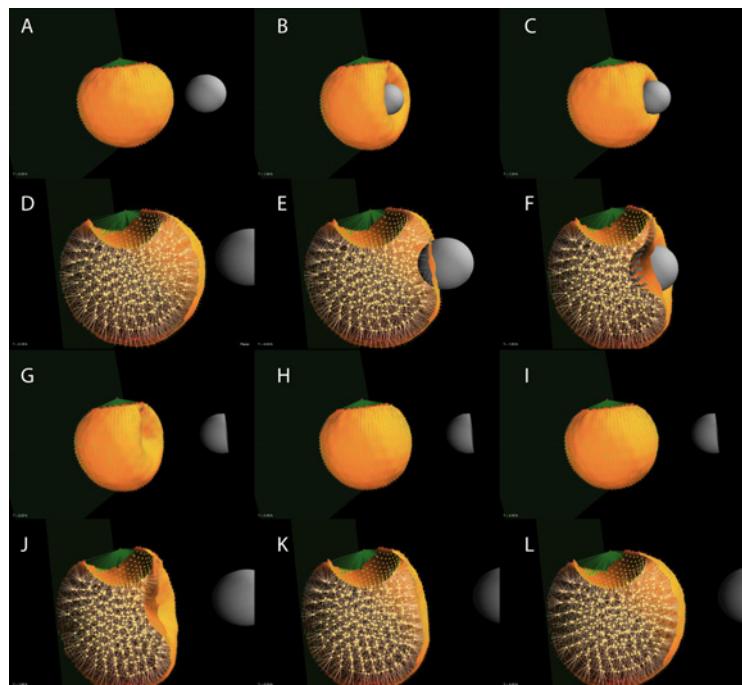


Figure 8.8: Snapshots of the simulated deformation and recovery. First and third rows: entire yolk from an external point of view. Second and fourth rows: another yolk sliced to reveal the yolk interior particles. A,B,C,D,E,F: compression phase. G,H,I,J,K,L: relaxation phase. Note that these simulated experiments have been realized by manually operating the spherical bead, whereas in the study the manipulation is automated. Original movies are available in Figs. S.3S.4

Fitness function

We define a fitness function $F_{Y,\text{rel}}$ which allows us to confront the simulations and the measurements: it is an integral of the evolution (here, a discrete sum over time) of the absolute difference between the live and the simulated normalized length, starting at time t_1 when the pressure is removed, and ending at $t_n = t_1 + (n - 1) \Delta t_y$ with typically $n = 10$. Denoting by $L^s(t)$ and $L^v(t)$ the deformed yolk diameters of the simulated and the live (“*in vitro*”) specimens measured at time step t , the equation of the fitness function reads:

$$F_{Y,\text{rel}} = \sum_{t=t_1}^{t_n} \left| \frac{L^s(t)}{L_{\text{rel}}^s} - \frac{L^v(t)}{L_{\text{rel}}^v} \right|$$

where L_{rel}^s and L_{rel}^v are the initial uncompressed yolk diameters. Note that the lower the fitness values, the better the match.

Fitness landscape

We explore the parameter space of the simulation to find the best match with the measurements. This space is 4-dimensional, comprising \bar{w}_{adh}^y , \bar{w}_{rep}^y , \bar{k}_{my} , and c_r , and was regularly sampled over the following ranges with the following numbers of values (“cardinalities”):

	Min.	Max.	Cardinality	Unit
\bar{w}_{adh}^y	0	$3.7 \cdot 10^{-5}$	4	s^{-1}
\bar{w}_{rep}^y	$3.7 \cdot 10^{-6}$	$5.55 \cdot 10^{-5}$	10	s^{-1}
\bar{k}_{my}	$3.7 \cdot 10^{-4}$	$3.7 \cdot 10^{-3}$	10	$\text{m}^2 \cdot \text{s}^{-1}$
c_r	0.5	1.0	10	

Figure 8.9: Ranges, cardinalities and units of the four parameters explored in this study.

For each “realization”, i.e. simulation under a particular set of parameters, the fitness is evaluated. The simulation time step Δt_s has been calibrated to match the slope of the relaxation curve obtained in the live experiment (Fig. 8.10) and its calibrated value is here $\Delta t_s = 6.66667$ milliseconds. This means that the yolk measurement time step Δt_y of 0.8 seconds is equivalent to 120 simulated time steps.

This plot indicates that the simulated phenotype leads to a relaxation profile similar to the profile provided by the live experiment. However, another representation was needed to visualize the fitness pattern in the 4D parameter space. We developed an interface displaying the fitness as a 3D cube and selecting the fourth dimension manually (Figs. 8.11, 8.12, and 8.13).

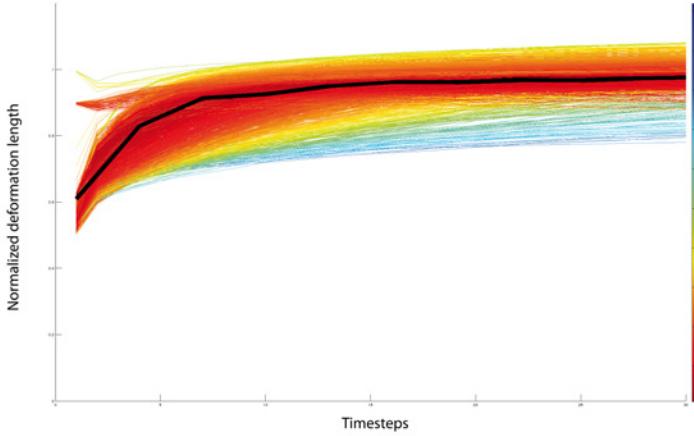


Figure 8.10: Plots of the simulated normalized deformation length as a function of time. The black thick curve represents the experimental live measurement. The simulated measurements are colored according to their fitness value (from red for the best fitness $F = 0.1609$ to blue for the worst $F = 7.1449$). In ordinate, the normalized length of the yolk diameter is plotted. In abscissa, the time from the moment the bead is pulled away from the yolk surface.

Fitness pattern analysis

We can observe patterns of low (i.e. optimal) fitness values in the left column. The shape of the isofitness patterns shows no changes along the depth axis, indicating that the adhesion coefficient between yolk interior particles does not impact the relaxation behavior of the yolk. Therefore, we can interpret these results by saying that it is the yolk membrane properties (stiffness and surface equilibrium) that account for the behavior of the system.

However, low values of yolk interior particle repulsion do induce a shrinking of the total yolk volume, as it can be seen in rows 1 to 5 of Fig. 8.12. Conversely, as the yolk interior repulsion increases, the total yolk volume increases. From a biological point of view this feature does not make much sense as we would rather expect the yolk to be incompressible, thus it rather points to one of the limits of our model. For this reason, we decided to limit the exploration of the parameter space to repulsion coefficient values that keep the yolk volume within a 20 percent variation of the targeted volume (right column).

In addition, when the yolk interior repulsion increases and the membrane stiffness is too high for a small yolk equilibrium surface, the numerical simulation of the model diverges and the phenotype can no longer be computed. This creates more and more missing points in the left and central column from the top to the bottom

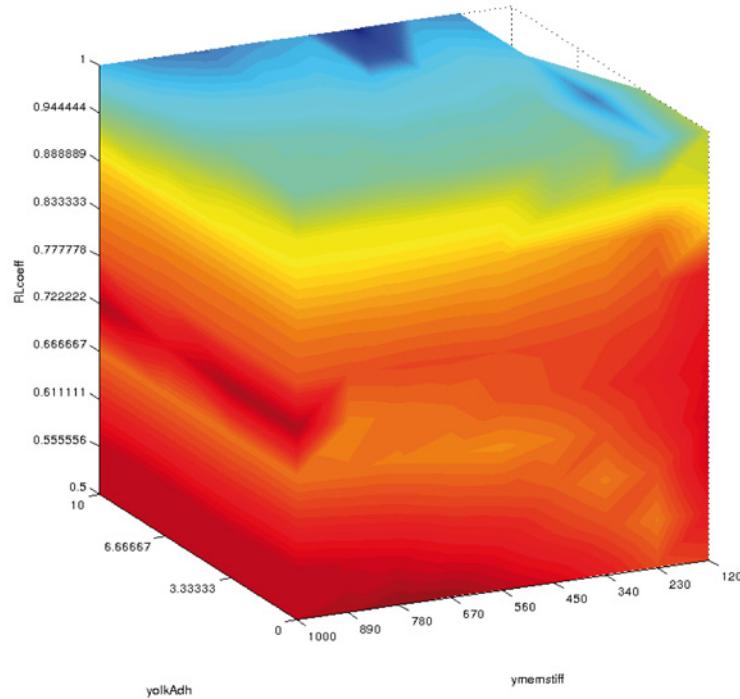


Figure 8.11: An example of 3D plot of the fitness landscape in parameter space. This figure aims at presenting the coordinate system used in Fig. 8.12 and Fig. 8.13. Here, the repulsion coefficient $\bar{w}_{\text{rep}}^{\text{y}}$ of the force exerted between yolk interior particles has a fixed value $3.7 \cdot 10^{-6}$. The vertical axis is the coefficient c_r , which determines the equilibrium surface of the yolk membrane. The depth axis is the adhesion coefficient, $\bar{w}_{\text{adh}}^{\text{y}}$ which modulates the cohesion of the yolk interior particles. The width axis is the yolk membrane stiffness coefficient \bar{k}_{ym} . It should be noted that the values plotted on the depth (resp. width) axis actually correspond to $w_{\text{adh}}^{\text{y}}$ (resp. k_{ym}) and should be divided by λ_0 to match the values of the table presented in Fig. 8.9.

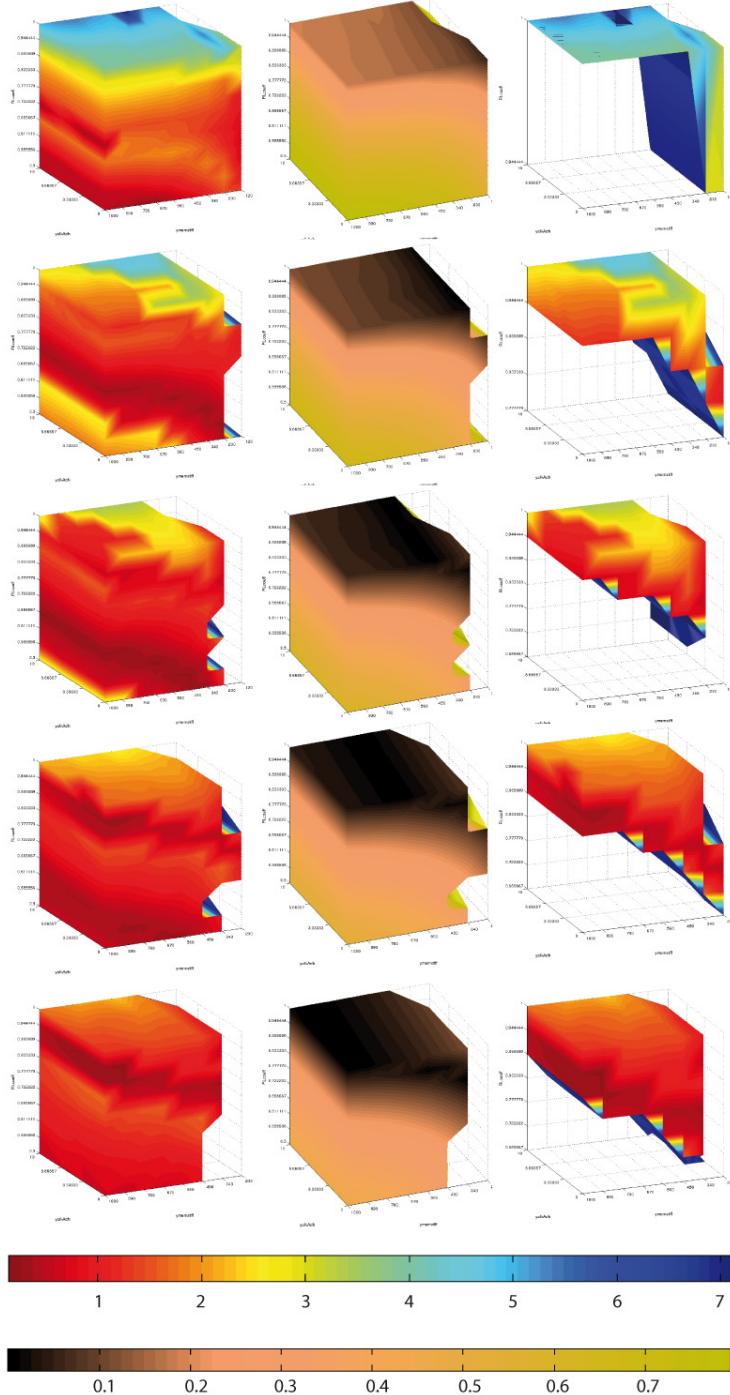


Figure 8.12: Fitness landscape of the yolk relaxation experiment. The left column represents the fitness landscape in parameter space. The middle column shows the normalized yolk volume difference between the simulated phenotype and the targeted volume of radius R_Y . The right column is identical to the left one but keeping only the phenotypes that have a yolk volume less than 20 percent different from the targeted one. All the coordinates are described in Fig. 8.11. Each row corresponds to a different value of the coefficient of repulsion between yolk interior particles \bar{w}_{rep}^y . The value of \bar{w}_{rep}^y varies regularly between $3.7 \cdot 10^{-6}$ for the top row to $5.55 \cdot 10^{-5}$ for the bottom row.

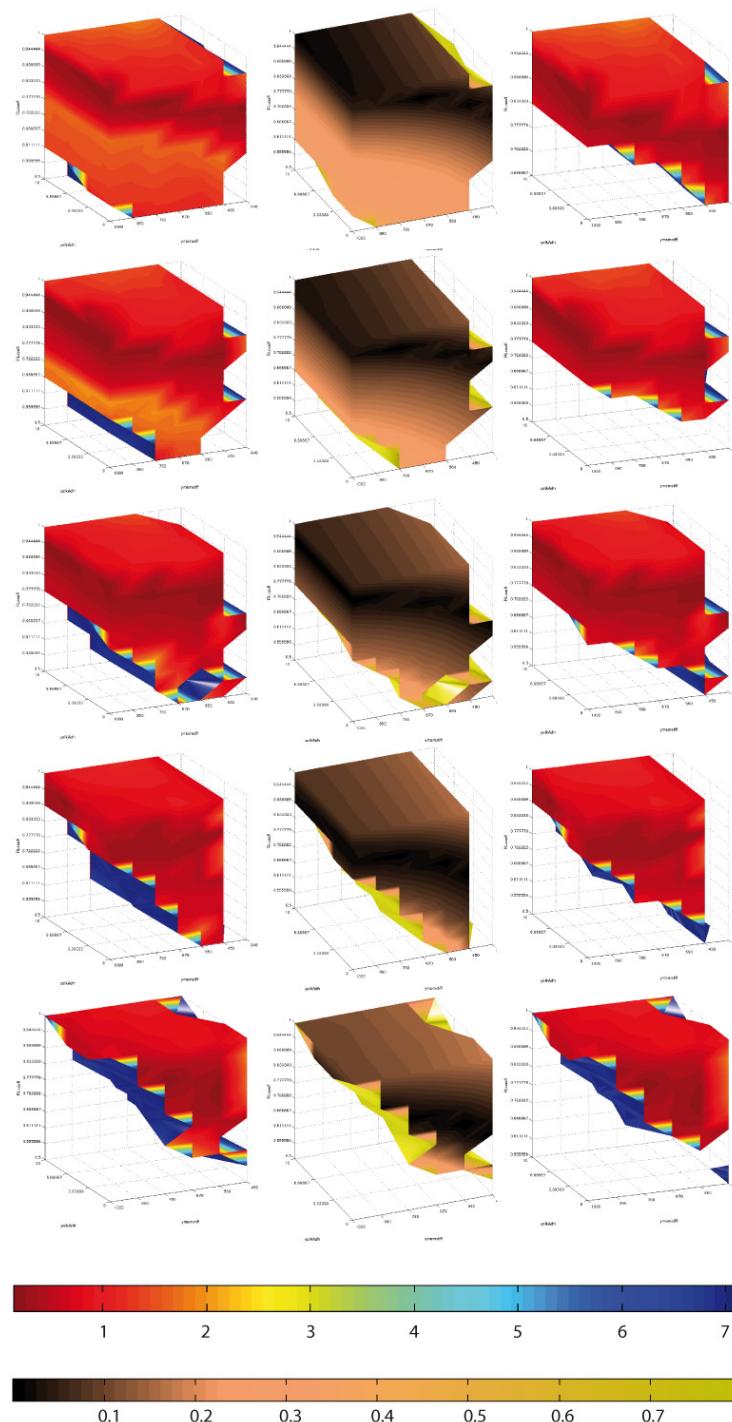


Figure 8.13: Fitness landscape of the yolk relaxation experiment. For caption, see Fig. 8.12.

rows. The “useful” space displayed in the right column indicate a single layer of low fitness value and the spatial orientation of this layer shows the benefits of increasing membrane stiffness *and* membrane surface equilibrium.

Therefore, the model suggests that the yolk membrane is responsible for the yolk relaxation behavior, the yolk interior particles only filling the space and maintaining the yolk volume. Again, “yolk membrane” here refers to membrane *and* cortex together, thus from a biological point of view the low fitness domain provides indications of appropriate surface equilibrium values for the yolk membrane + cortex ensemble.

8.1.3 Discussion

This very simple case study already illustrates both the potential and the limitations of our model, raising a number of generic questions. The choice of an agent-based framework led to the representation of the yolk and yolk membrane as collections of interacting particles. If this type of model seems indeed appropriate to mimic some of the observed properties of the biological system under study, the effect of yolk interior particles’ repulsion and attraction coefficients on the yolk volume is unlikely to be very biologically relevant. We would rather think of the yolk interior as an incompressible liquid, although this hypothesis has not been validated experimentally either. In any case, large variations created by simulation must always be discarded as artifacts. Still, such shortcomings do not prevent exploring the membrane properties within a reasonable range of values and try to give a biological interpretation to the resulting fitness landscape.

The numerical implementation of the model contains a number of difficulties, including instabilities of the numerical integration that prevented to complete the calculation of certain phenotypes. It could still be said, however, that most of the parameter space points that were not explored were actually unlikely to correspond to the best fitness pattern. In addition, the computation time step that had to be set to fit the experimental scenario of the yolk relaxation is far too small, hence computationally expensive, for the simulation of embryo development until the end of gastrulation, i.e. 10 hours post fertilization (hpf).

Interpreting the calculated fitness landscape first raised the problem of the representation and visualization of high-dimensional data. Furthermore, we did not fully explore the parameter space here. Some of the observed behaviors suggest that several isofitness layers are created by exploring the yolk interior’s repulsion values. From row 4 in the left column of Fig. ??, the coexistence of two low isofitness layers suggests that a periodic behavior might happen. Further exploration of values along the vertical axis should be done. Hence, we so far have no conclusive explanation of this behavior of the model. In general, the biological and quantitative interpretation of this model is an issue. We expect at some point to be able to derive realistic physical quantities from the simulation parameters. Since the simulation

values are scaled by the viscosity factor λ_0 , additional stress-strain measurements would be needed to tune this parameter other than arbitrarily.

In conclusion, this simple experimental protocol should open the way to a more extensive study that would (i) provide measurements for a cohort of individuals, (ii) compare the yolk behavior in different locations at different developmental times, (iii) control the applied force and possibly replace the pushing strategy by pulling with a known aspiration force. In addition, the measurement of the deformation might be more precise and accurate than accomplished here. Finally, in addition to designing a fitness function and finding its optimal (i.e. lowest) values, the model should also propose a definition for a domain of *viability*. The fitness landscape and domain of viability might then be adjusted together by exploring other desirable properties of the yolk compatible with later morphogenetic events, such as doming or the typical constriction accompanying gastrulation movements.

8.2 Cell Proliferation Rate Along the Cell Lineage

Cell proliferation increases the number of embryonic cells and is the basis for cell diversification in time and space. The proliferation rate of embryonic cells is characteristic of their potentialities and is subjected to regulation in time, then in space, leading to different mitotic domains and, at later embryonic and larval stages, to cell populations with very different proliferation rates.

Our model should be useful to test different theoretical rules modulating the cell proliferation rate along the cell lineage. However, except for the cells of the enveloping layer (EVL; see Section 8.4), the cell cycle in our model is uncoupled from other cellular mechanisms and in particular from biomechanical properties. This means that we are not making any hypothesis on the possible correlation between the cell cycle length and morphogenetic events. However, we should be able to investigate the possible correlation of cycle length with cell volume and test simple lengthening rules, such as a geometric progression, which were proposed by state-of-the-art studies [309].

Proliferation during cleavage stages is under maternal control. Proliferation along the cell lineage has been measured precisely for the first 10 cell cycles in early zebrafish embryos by Olivier et al. [254] (Fig. 8.14).

This quantitative study suggests a single dynamic regime for this developmental period, during which the cell cycle length is first shortening then lengthening, although this variation in duration remains quite small (no more than a few minutes). Cell cycle lengthening is supposed to reflect the exhaustion of maternal factors driving the cell cycle progression. It has been postulated that the lengthening of the cell cycle during blastula stages correlates with the ratio of the nucleus size over the cell size. This correlation, however, would not exist any longer by cell cycle 14 [280]. In Olivier et al. [254], no correlation was found between cycle duration and

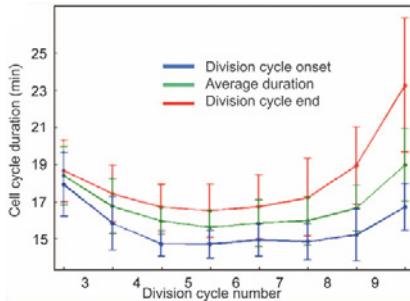


Figure 8.14: Mean cell cycle duration (green) compared with cycle duration at the beginning (blue) and the end (red) of the division cycle, averaged over six embryos. Image and caption adapted from [254].

cell volume for the first 10 cell cycles. Rather, cell cycle length was observed to be dependent on the cell position within the blastoderm. That study proposed that division asynchrony is the rule and increases from cycle 2 to cycle 10 (Fig. 8.15).

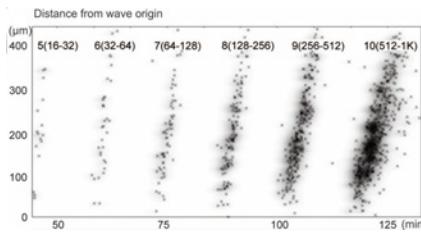


Figure 8.15: Scatter plot of mitosis time as a function of distance to the pseudo-wave origin. Image and caption adapted from [254].

This conclusion differs somewhat from the work by Kane and Kimmel [309] indicating a measurable asynchrony by the midblastula transition only (Fig. 8.16). This discrepancy might however solely reflect the measurements method's sensitivity and precision.

Measurements plotted in Fig. 8.16 highlight the progression of cell cycle duration throughout the midblastula transition starting at cycle 10 (512 to 1024 cells, 3 hpf). The cell cycle length increases very significantly as soon the zygotic transcription begins and throughout cycles 11 and 12 [309]. Cycle 12 should be completed around the sphere stage (4 hpf). According to Kane and Kimmel [309], cell cycle lengthening would respond to a single dynamical regime during blastula and early gastrula stages, with the cell cycle length progressing as a linear function of 1/cell volume (Fig. 8.17). Such a correlation would lead to a geometric progression of cell cycle lengthening, doubling at each cycle. It should however be noted that this rule does fit with the measurements provided in the paper.

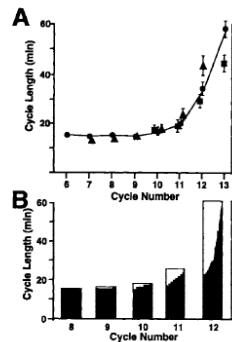
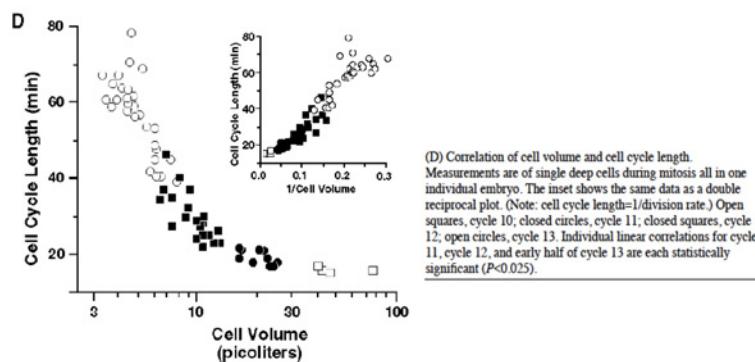


Fig. 9. Cell cycle lengthening and loss of synchrony at the midblastula transition, beginning at the tenth zygotic cell cycle (512-cell stage). The three different symbols in A represent data collected by following fields of cells in three different embryos. The distributions in B represent cycle lengths of individual cells followed in a field within a single embryo. Reproduced from Kane and Kimmel (1993) with permission from Company of Biologists Ltd.

Figure 8.16: Cell cycle lengthening and loss of synchrony at the midblastula transition. Adapted from [309].



(D) Correlation of cell volume and cell cycle length.
Measurements are of single deep cells during mitosis all in one individual embryo. The inset shows the same data as a double reciprocal plot. (Note: cell cycle length=1/division rate.) Open squares, cycle 10; closed circles, cycle 11; closed squares, cycle 12; open circles, cycle 13. Individual linear correlations for cycle 11, cycle 12, and early half of cycle 13 are each statistically significant ($P < 0.025$).

Figure 8.17: Correlation of cell volume and cell cycle length for cycles 10 to 13. Adapted from [309].

According to Kane and Kimmel, gastrulation would proceed during cell cycle 14. Cell proliferation during gastrulation continues to slow down dramatically so that, on average, cells divide only once between 6 and 10 hpf. It has also been shown that gastrulation is not impaired by the inhibition of cell divisions. If cell proliferation is indeed required to provide a minimal number of cells for further morphogenetic movements, gastrulation appears quite robust to perturbations of the cell cycle. For example, in *emi1* (*early mitotic inhibitor-1*), homozygous mutant defective in the expression of this negative regulator of the Anaphase Promoting Complex [310], gastrulation and somitogenesis proceed while mitotic activity stopped at the beginning of gastrulation (shield stage 6 hpf).

The state-of-the-art data and analysis leave open a number of questions. First, precise measurements of the cell cycle duration along the cell lineage and other parameters including the cell volume and nucleus volume, when discussing correlations between the cell cycle length and the cell volume or nucleo-cytoplasmic ratio, are lacking except for the first 10 cycles [254]. Without precise phenomenological reconstructions, it is very difficult to make relevant hypotheses on the dynamic regimes controlling the cell proliferation rate along the cell lineage. The model proposed by Kane and Kimmel for the progression of the cell cycle duration in $1/V$ readily leads to a geometric progression rule that should be further tested and discussed. In addition, increasing asynchrony of cell divisions should be better characterized to provide the most precise phenomenology for a further investigation of the underlying mechanisms.

8.2.1 Hypotheses and Model

In our model, we make the approximation that through the first 10 divisions, cell cycle length is constant and constrained to reach the 1,000-cell stage by 3 hpf, i.e. at the time of midblastula transition. We are then looking for an appropriate model of cell cycle lengthening and subsequent desynchronization. Our approximation does not exclude, however, that the desynchronization might obey the same rule throughout the first 14 or 15 cell cycles. From the study by Kane and Kimmel [309], we understand that the cell cycle duration might slow down according to a geometric progression. Indeed, the authors suggest a progression as a linear function of $1/V$. As cells are expected to divide at constant volume until the end of gastrulation, the cell volume should be divided and the cell cycle should double. Yet, as already discussed, this does not match the data (Fig. 8.18).

We propose to test two different hypotheses corresponding to either a geometric or an arithmetic progression of cell cycle lengthening. In each model, two parameters determine the cell cycle progression: one for the mean of the distribution and the other for its range. In the following, the length of the cell cycle is denoted by L_{cc} . It is function of an integer value n representing its iteration, e.g. $L_{cc}(12)$ is the length of the 12th cell cycle. The iteration from which the cell cycle increases is denoted by

Table 1. Median cell cycle lengths during zebrafish midblastula stages at 28°C

Cycle number	Median cycle length (min)	90% confidence limits*
7	14.2	12.9-15.0 (n=9**)
8	14.5	13.8-15.5 (n=13)
9	15.0	14.5-15.9 (n=16)
10	17.0	16.1-18.0 (n=17)
11	22.5	20.5-27.8 (n=17)
12	33.5	31.6-36.7 (n=7)

*Nonparametric confidence limits calculated from the Sign Test (Fisher, 1925; Thompson, 1936).

**Number of embryos analysed for each cycle number.

Figure 8.18: Median cell cycle lengths during blastula stages. Adapted from [309].

n_{start} and the last iteration by n_{end} . Similarly, the timing of a cell cycle is denoted by T_{cc} .

Geometric progression of the cell cycle

A geometric progression of the cell cycle iteratively multiplies the length of the cell cycle by a value, the *common ratio*, denoted by r :

$$\forall n \in [n_{\text{start}}, n_{\text{end}}], L_{\text{cc}}(n+1) = r \cdot L_{\text{cc}}(n)$$

We suppose that the common ratio follows a uniform distribution centered on r_0 and with a range of w_r :

$$r = \mathcal{U}(r_0 - \frac{w_r}{2}, r_0 + \frac{w_r}{2})$$

During the development of a simulated embryo, each cell division requires the generation of a new random common ratio r determined with the parameters r_0 and w_r . We also make the assumption that the cell cycle length is always increasing after each division, so the common ratio r must be larger than 1. It implies the following constraint on the parameters of the study:

$$w_r \leq 2(r_0 - 1)$$

Arithmetic progression of the cell cycle

An arithmetic progression of the cell cycle iteratively sums the length of the cell cycle by a constant value, the *common difference*, denoted by d :

$$\forall n \in [n_{\text{start}}, n_{\text{end}}], L_{\text{cc}}(n+1) = d + L_{\text{cc}}(n)$$

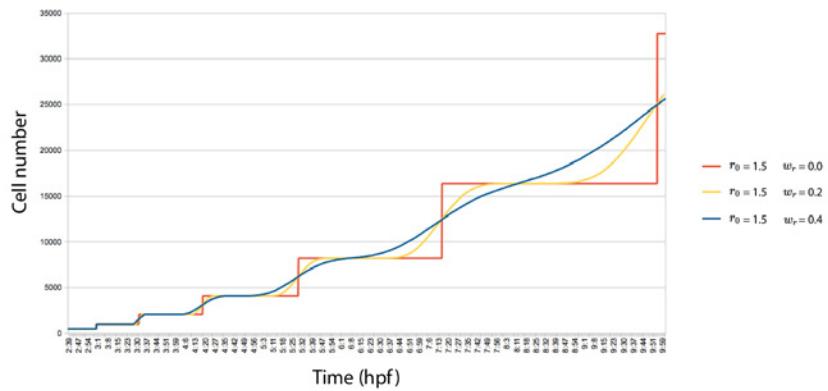


Figure 8.19: Theoretical cell number as a function of time of development in the geometric progression scenario. In ordinate, the number of simulated cells. In abscissa, time of development starting at 2h39. Three geometric progression scenario are plotted. Each of them starts from an initial simulated cell population comprising 512 cells with identical cell cycle length (see text), and identical common ratio ($r_0 = 1.5$). In red, the asynchrony is zero, all cells divide simultaneously at each cycle ($w_r = 0$). In yellow, cell cycles are desynchronizing reasonably at each cycle ($w_r = 0.2$), plateaus are still occurring in between mitosis cycles. In blue, a larger desynchronization ($w_r = 0.4$) leads to a progressive disappearance of the plateau as the cell population increases.

As before, we assume that the common difference follows a uniform distribution centered on d_0 and with a range of w_d .

$$d = \mathcal{U}(d_0 - \frac{w_d}{2}, d_0 + \frac{w_d}{2})$$

During the development of a simulated embryo, each cell division requires the generation of a new random common difference d determined with the parameters d_0 and w_d . We also make the assumption that the cell cycle length is always increasing after each division, so the common difference d must be larger than 0. It implies the following constraint on the parameters of the study:

$$w_d \leq 2r_0$$

8.2.2 Simulation, Parameter Space and Validation

Cell cycle length measurement through the reconstructed lineage tree

To explore the parameter space and evaluate the fitness of the two hypothesized cell cycle rules (geometric and arithmetic), we manually selected a small population of cells from cell cycle 9 (512-cell stage) at approximately 2h40 hpf and followed them through the lineage tree produced by the reconstruction workflow (described in Chapter 7) until cell cycle 15.

As we do not have the complete data for cell cycle 15, this data is not included in the study. In addition, we interpret the three points that are outside the clouds to be caused by tracking errors that were not found by eye inspection of the raw data. These three points have also been removed from the study.

Simulated cell cycles

We simulated two developments with respectively a geometric progression and an arithmetic progression of the cell cycle length between cycle 9 and cycle 14. To account for the desynchronization of the cell cycles, we took the common ratio and common difference as uniform distributions, as described above. We derived from these simulations the same kind of plot as for the live experimental data (Fig. 8.23). However, unlike for the live experiment, we were able here to plot all the cells of the embryo.

Fitness function

The fitness function used to evaluate the adequacy of the proposed cell cycles is multi-objective. Four objective subfunctions will be evaluated for each set of parameters:

- $F_\mu(T_{cc})$, the mean of the cell cycle timing for each generation,

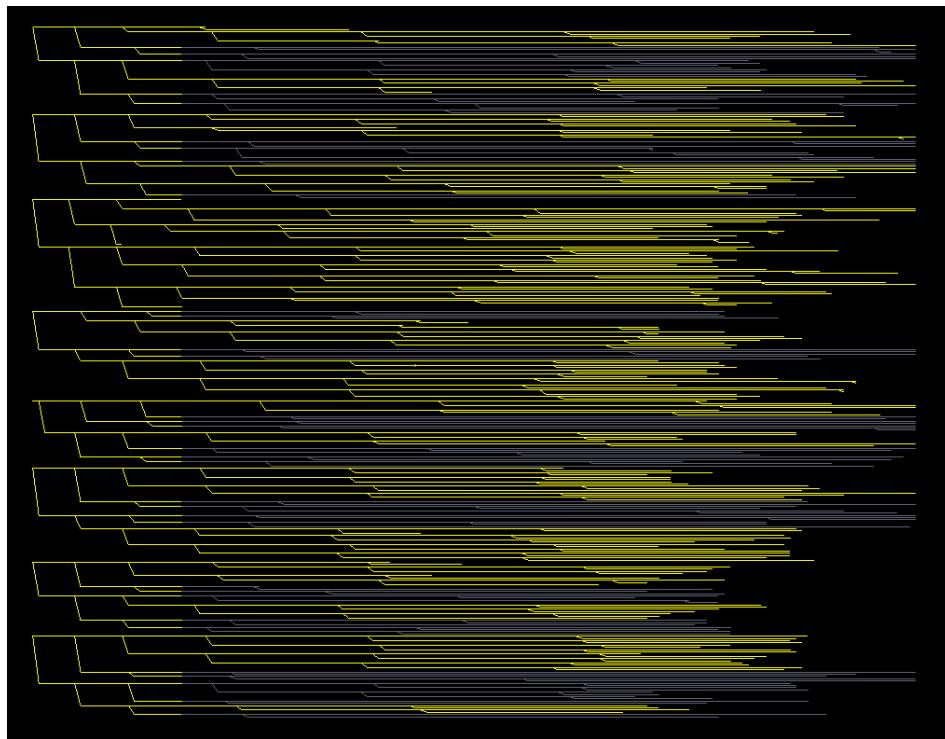


Figure 8.20: Lineage tree of the selected cell population.
The selected population consists in eight cells picked at the time of cell cycle 9. At cycle 9 they divide quasi-synchronously (one cell out of the eight divides one time step, i.e. 3 min, after the others). The selected cells contribute to the formation of the EVL. As this differentiated tissue is expected to have a specific cell cycle dynamics (see Case Study 8.4), EVL cells were removed from the selected cell population after division cycle 11. EVL cell lineage removed from the analyzed data is indicated by the grey branches. Cell cycle 11 occurs at the onset of the epiboly at the time of doming, prior to what is usually considered as gastrulation (6 hpf). The selected cells tracking was checked and corrected as far as possible in time.

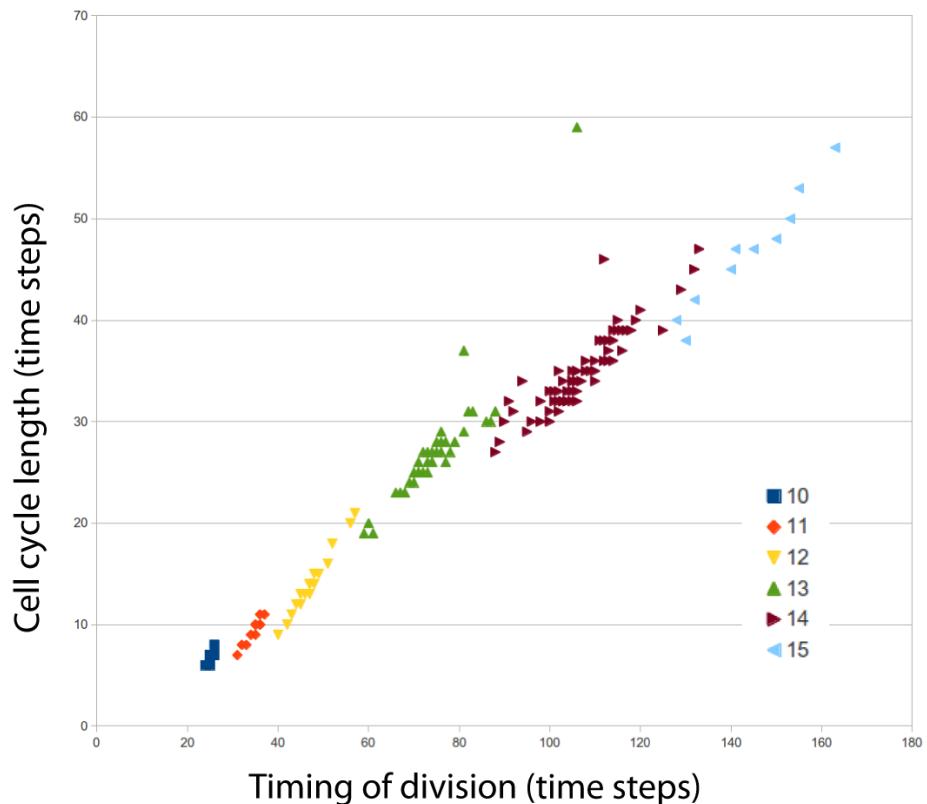


Figure 8.21: Plot of the cell cycle length for cells selected among the deep cells, as a function of the timing of the divisions. Colors and numbers represent the iteration of the cell cycle as indicated on the right. A cell is plotted at the time of a division as a function of the timing of its mother's division. Timing is indicated in time steps. Data manually validated by Ingrid Colin with Mov-IT.

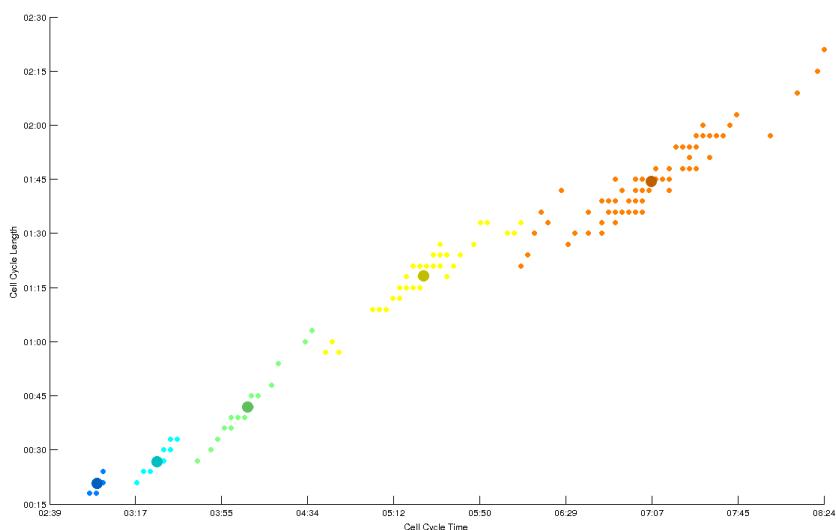


Figure 8.22: Plot of the cell cycle length as a function of the cell cycle time for the cell population selected in the digital specimen. The blue color represents the 10th cell cycle, light blue, the 11th, green, the 12th, yellow, the 13th, and orange, the 15th. Small dots correspond to single cells at the time of mitosis and large dots are the mean of the division timing for the corresponding cell cycle. The data is the same as Fig. 8.21 without cell cycle 15 and after elimination of the three outliers discussed above. This data is used for confronting live experimental and simulated data.

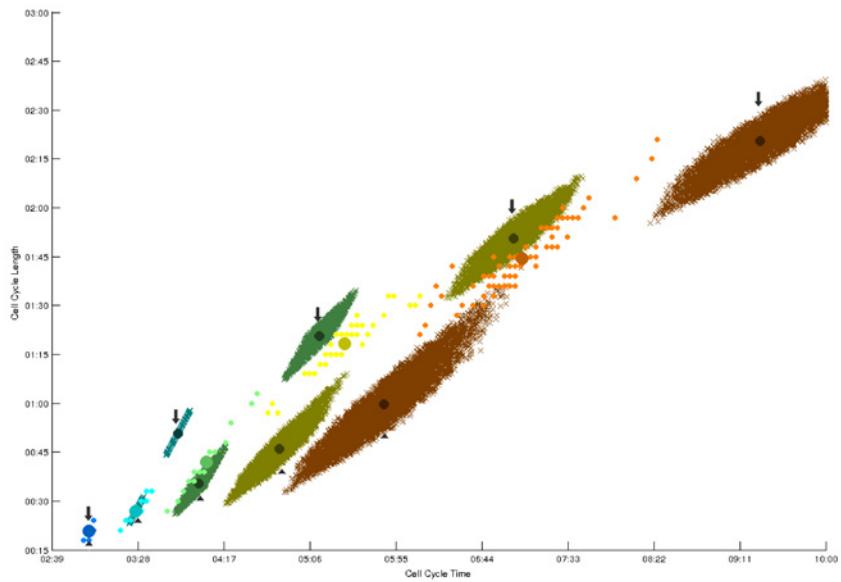


Figure 8.23: Plot of the cell cycle length as a function of the the cell cycle time for two simulated embryonic developments. In addition to the live cells plot, we provide two simulated distributions, one corresponding to a geometric progression, indicated by arrowheads ($r_0 = 1.3$ and $w_r = 0.4$, and one corresponding to an arithmetic progression, indicated by arrows ($d_0 = 30\text{min}$ and $w_d = 14\text{min}$). For the sake of visualization, we picked here parameters that obviously did not provide the best fitness but kept the clouds separated.

- $F_\sigma(T_{cc})$, the standard deviation of the cell cycle timing for each generation,
- $F_\mu(L_{cc})$, the mean of the cell cycle length for each generation,
- $F_\sigma(L_{cc})$, the standard deviation of the cell cycle length for each generation.

All four objectives are expressed through the same type of generic objective function:

$$\begin{aligned} F_\mu(T_{cc}) &= \sum_{n=n_{\text{start}}}^{n_{\text{end}}} \frac{|\mu[T_{cc}^s(n)] - \mu[T_{cc}^v(n)]|}{\mu[T_{cc}^v(n)]} \\ F_\sigma(T_{cc}) &= \sum_{n=n_{\text{start}}}^{n_{\text{end}}} \frac{|\sigma[T_{cc}^s(n)] - \sigma[T_{cc}^v(n)]|}{\sigma[T_{cc}^v(n)]} \\ F_\mu(L_{cc}) &= \sum_{n=n_{\text{start}}}^{n_{\text{end}}} \frac{|\mu[L_{cc}^s(n)] - \mu[L_{cc}^v(n)]|}{\mu[L_{cc}^v(n)]} \\ F_\sigma(L_{cc}) &= \sum_{n=n_{\text{start}}}^{n_{\text{end}}} \frac{|\sigma[L_{cc}^s(n)] - \sigma[L_{cc}^v(n)]|}{\sigma[L_{cc}^v(n)]} \end{aligned}$$

where $n_{\text{start}} = 10$, $n_{\text{end}} = 14$, $L_{cc}^v(n)$ (resp. $T_{cc}^v(n)$) is the ensemble of cell cycle timing (resp. length) measured on the real embryo at cell cycle n , $L_{cc}^s(n)$ (resp. $T_{cc}^s(n)$) is the ensemble of cell cycle timing (resp. length) measured on the simulated embryo at cell cycle n , μ is the mean operator and σ the standard deviation operator. Each absolute difference is normalized by the value of the experimental measure, such that the contribution of each cell cycle generation's mean or standard deviation has the same influence on the objective function.

These four objective functions are then merged into a global fitness function through a weighted sum method. Applied to the parameter space, this generates four objective landscapes for each proposed model (Fig. 8.25 and Fig. 8.26). As the aim of this study is to compare the two hypothesized cell cycle rules (geometric and arithmetic) through a global fitness function, each landscape value is normalized to contribute equally to the global fitness function. Various methods of normalization can be used [311]. Here, we simply divide each objective function by the maximum objective function value over both cell cycle rule landscapes:

$$\overline{F} = \frac{F}{\max(F^{\text{geo}}, F^{\text{ari}})}$$

where F stands for any of the above four objective functions. The global fitness F_{cc} function is then the average of the four normalized objective functions:

$$F_{cc} = \frac{1}{4} (\overline{F_\mu}(T_{cc}) + \overline{F_\sigma}(T_{cc}) + \overline{F_\mu}(L_{cc}) + \overline{F_\sigma}(L_{cc}))$$

	Min.	Max.	Cardinality	Unit
r_0	1	2	21	-
w_r	0	2	41	-
d_0	0	40	21	min
w_d	0	80	41	min

Figure 8.24: Ranges, cardinalities and units of the four parameters explored in this study.

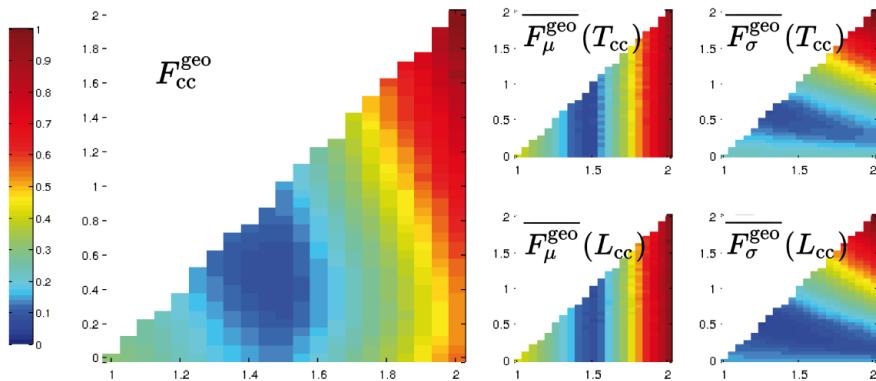


Figure 8.25: Fitness landscapes of the *geometric progression rule* in parameter space. The left plot is the global fitness landscape. In the right part of the figure: the top left plot is the landscape of the mean of the cell cycle time's normalized objective function $\overline{F}_\mu^{\text{geo}}(T_{cc})$, the top right plot is the landscape of the standard deviation of the cell cycle time's normalized objective function $\overline{F}_\sigma^{\text{geo}}(T_{cc})$, the bottom left plot is the landscape of the mean of the cell cycle length normalized objective function $\overline{F}_\mu^{\text{geo}}(L_{cc})$, the bottom right plot is the landscape of the standard deviation of the cell cycle length normalized objective function $\overline{F}_\sigma^{\text{geo}}(L_{cc})$. The color map of all five plots follows the same color bar presented on the left. Each plot provides the cell cycle distribution center r_0 on the abscissa axis and the distribution range coefficient w_r on the ordinate axis. Both parameters are explored in the range presented in Fig. 8.24

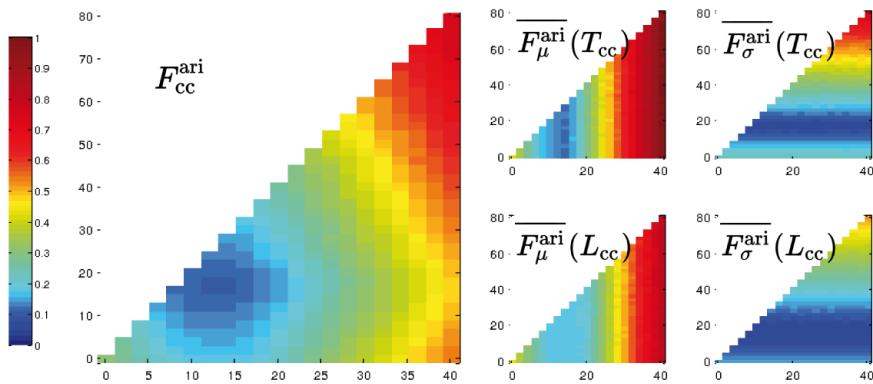


Figure 8.26: Fitness landscapes of the *arithmetic progression rule* in parameter space. The left plot is the global fitness landscape. In the right part of the figure: the top left plot is the landscape of the mean of the cell cycle time's normalized objective function $\overline{F}_\mu^{\text{ari}}(T_{cc})$, the top right plot is the landscape of the standard deviation of the cell cycle time's normalized objective function $\overline{F}_\sigma^{\text{ari}}(T_{cc})$, the bottom left plot is the landscape of the mean of the cell cycle length normalized objective function $\overline{F}_\mu^{\text{ari}}(L_{cc})$, the bottom right plot is the landscape of the standard deviation of the cell cycle length normalized objective function $\overline{F}_\sigma^{\text{ari}}(L_{cc})$. The color map of all five plots follows the same color bar presented on the left. Each plot provides the cell cycle distribution center d_0 on the abscissa axis and the distribution range coefficient w_d on the ordinate axis. Both parameters are explored in the range presented in Fig. 8.24

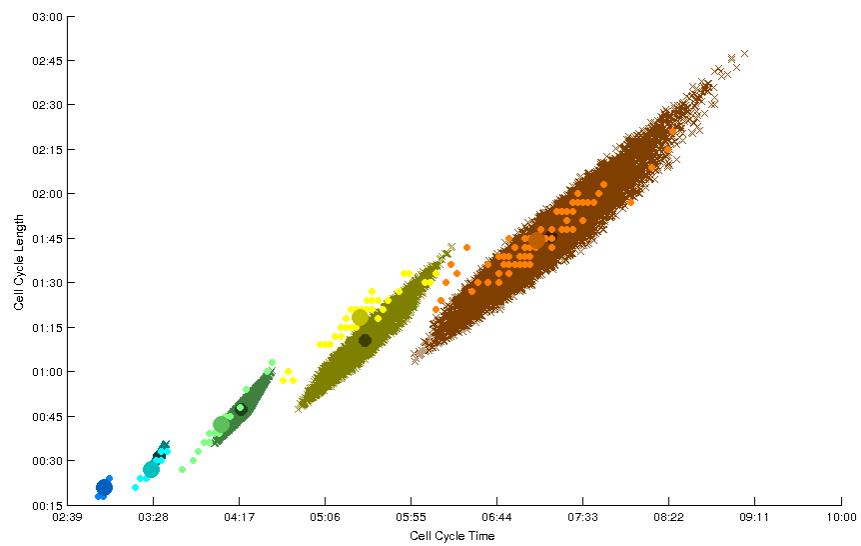


Figure 8.27: Plot of the cell cycle length L_{cc} as a function of the cell cycle time T_{cc} for the simulated embryonic development with the best set of parameters for the *geometric progression rule*. This phenotype has also the best fitness of this study ($r_0 = 1.5$, $w_r = 0.4$).

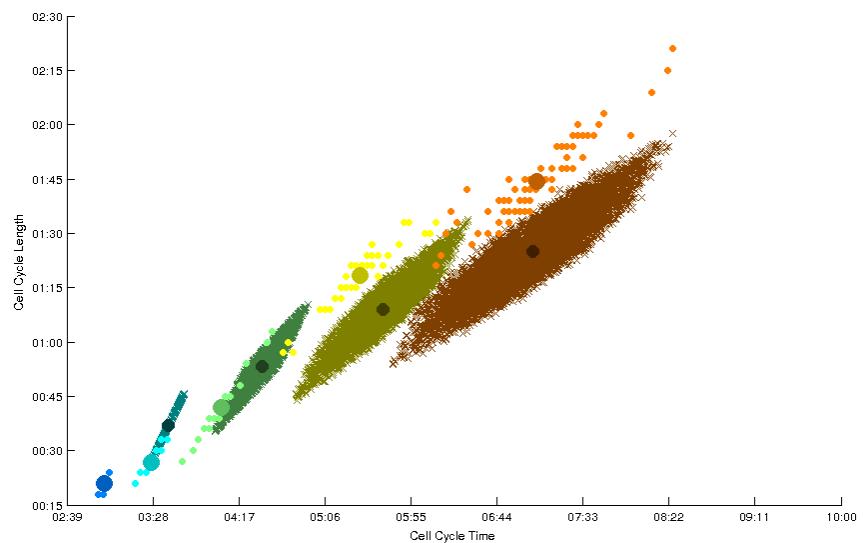


Figure 8.28: Plot of the cell cycle length L_{cc} as a function of the cell cycle time T_{cc} for the simulated cells generated with the best set of parameters for the *arithmetic progression* rule. This phenotype corresponds to the following parameters: $d_0 = 14$ min, $w_d = 18$ min.

Fitness landscape

8.2.3 Discussion

Although the best fitness is realized by the geometric progression, the differences between the real and the simulated phenotypes derived from the two rules are rather close. These values would become more clearly distinct with a larger number of cycles, but this would not be very meaningful biologically, as the regulation of cell proliferation is expected to diversify at later stages and become dependent on cell location and/or lineage.

The best fitness for the arithmetic progression rule does not perform well for the mean time of each cell generation (Fig. 8.26 bottom left small plot). Thus the overall performance could have been different, i.e. resulting in a slight advantage for the geometric progression model, if the mean objective functions had been assigned a higher weight than the standard deviation functions in the global fitness expression F_{cc} .

Interestingly, the hypothesis of a geometric progression was proposed by Kane. However, this author hypothesized a progression function of $1/V$, meaning a common ratio of 2, assuming that at each cell division the cell volume is divided by 2 until the end of gastrulation. Kane's experimental data was not consistent with a common ratio of 2, however, and our best fitness suggests a common ratio of 1.5 instead. If the global cell volume conservation hypothesis was not correct, we could reconcile the two studies and calculate the expected volume evolution.

To account for the desynchronization of cell divisions, the common ratio and common difference for the geometric and arithmetic progressions were assumed to obey uniform distributions. This was an ad hoc choice and the possibility that other distributions would provide better fitness should also be explored. In addition, we simplified the system by assuming the first 9 cycles synchronous. Although this might be considered a reasonable approximation, in fact consistent with our data since our 8 cells divided almost synchronously at cycle 9, this might also hide other important features of the system. The point is that our model is probably well suited to analyze the overall desynchronization dynamics, and decide whether different desynchronization regimes occur throughout the first 14 or 16 cell cycles or whether, on the contrary, there is unique regime and the same type of distribution is suitable to evaluate the dispersion of division times along the lineage. The measurements provided by Olivier et al., which indicate that asynchrony starts at cycle 2 and increases through successive cycles might be used to test this hypothesis.

8.3 Shaping the Zebrafish Blastula

In this third case study, our objective is to elucidate the biological mechanisms at the basis of cell-cell and cell-yolk interactions that drive the shaping process of the

zebrafish blastula. As shown in previous studies [254], by the 16-cell stage, the 4 central cells tend to divide more orthogonally to the yolk sphere's tangential plane, leading to a two-layer blastoderm. Through successive divisions, the blastoderm acquires an increasing height on top of the yolk until the “high stage”, at which point it seems to start flattening, eventually giving rise to the so-called “sphere stage” (Fig. 8.29).

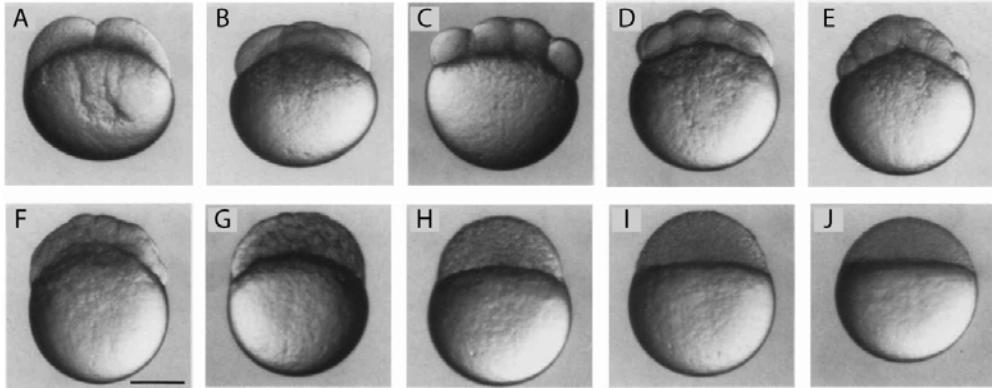


Figure 8.29: Cleavage and blastula period. A-F: Cleavage period. G-J: Blastula period. All face views except for B, which shows the embryo twisted about the animal-vegetal axis, roughly 45 degrees from the face view. A: 2-cell stage (0.75 h). B: 4-cell stage (1 h). C: 8-cell stage (1.25 h). D: 16-cell stage (1.5 h). E: 32-cell stage (1.75 h). F: 64-cell stage (2 h). G: 256-cell stage (2.5 h). H: high stage (3.3 h). I: transition between the high and oblong stages (3.5 h). J: transition between the oblong and sphere stages (3.8 h). Scale bar in F.: 250 μ m. Image and caption adapted from Kimmel et al. (1995) [253]

8.3.1 Hypotheses and Model

We hypothesize that until the early blastula stages (starting at the high stage), blastoderm shaping essentially arises from cell proliferation, the orientations of cell division, and cell-cell and cell-yolk attraction/repulsion forces with negligible intrinsic cell motility. We examined and discussed cell proliferation at the cleavage and blastula stages in the previous case study (Section 8.2). Cell division orientation has also been quantified in the literature [254]: at first orthogonal to each other, division planes appear to become more and more randomly oriented. However, we expect stronger biomechanical constraints on the blastoderm surface to correlate with a more biased orientation pattern. Division orientation in the outmost cell layer of the blastoderm is what distinguishes the EVL (the enveloping layer differentiating at the surface of the embryo to form an epithelium) from inner cell lineages. The EVL

starts differentiating as early as the 64-cell stage and becomes a closed compartment at the sphere stage (4 hpf) [257]. Between cell cycle 9 (256 to 512-cell stage) and cell cycle 10 (512 to 1K-stage), cell divisions at the blastoderm-yolk interface also lead to the formation of the yolk syncytial layer (YSL) [312].

In our model, blastula shaping results from the mechanical interactions between all the particles types involved in the processes described above. We distinguish here four different types of particles (Fig. 8.30):

- 1. The first category consists of the deep cells (DC, red circles in the 2D schema) whose interactions are driven by the generic rules described in Chapter 3.
- 2. Then, the yolk particles are separated into two categories: the yolk interior particles (YI, pale yellow triangles), and
- 3. the yolk membrane particles (YM, yellow squares), which form an viscoelastic sac surrounding YI. Interactions among these two types of yolk particles were described in the first case study (Section 8.1).
- 4. A new type of particle is introduced here: the EVL cell particles (blue crosses), which constitute the outmost layer of the blastoderm.

Following the template of the “passive adhesion table” $P = \{w_{TT}^{\text{adh}}\}$ of Section 5.2.3, mechanical interactions among particles are specified here for each pair of particle types (Fig. 8.30B). Note, however, that the present table is not composed purely of adhesion coefficients but rather “interaction types”, which might involve slightly different potentials and forces than the foundational model explained in Chapter 3. With the present goal of defining a relevant parameter space for the blastula shaping study, we will explain in the subsections below the four pairwise interaction types that have not been described yet (gray boxes in the table): YM-DC, EVL-DC, EVL-YM, and EVL-EVL. Note that two pairs of types do not correspond to any interaction as their cells are physically separated and never come into contact: YI-DC and EVL-YI. Thus the table contains a total of eight different interaction types.

Cleavage events and EVL formation in the model

The orientation of the first 4 cycles is constrained in the plane tangential to the yolk at the yolk-blastoderm interface (Fig 8.31, axes A_1 and A_2), leading to a 16-cell stage embryo with a single cell layer at the yolk surface (Movie 8.5). Moreover, during the first divisions of the cleavage period, cytoplasmic material is transferred from the yolk cell to the deep cells (see Fig. 8.2). We model this phenomenon by increasing the deep cells’ radii at each division until cell cycle 7 (not shown in Fig. 8.31; would correspond to letter H). The DC growth ratio is calculated to obtain

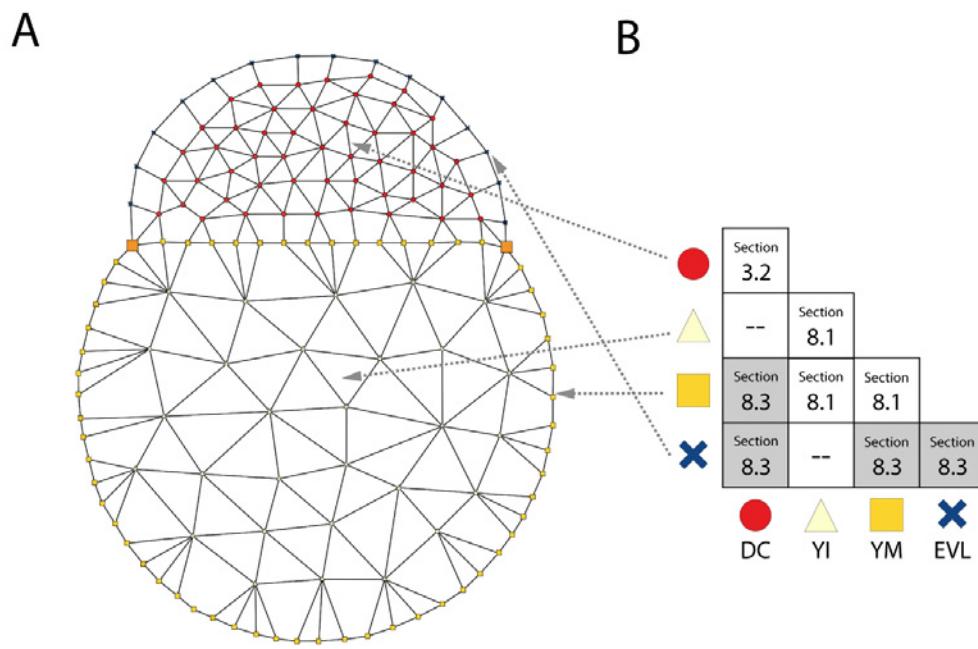


Figure 8.30: Complete abstract schema of the different particle types involved in the modeled embryo and their interactions. Red circles represent the deep cells (DC), yellow squares the yolk membrane (YM) particles, pale yellow triangles the yolk interior (YI) particles, and blue crosses the EVL cells. A: The two bigger orange squares represent YM particles that are labeled “margin” as the EVL is attached to the yolk in this position. B: Table of interactions by particles types. Each box indicates the section of the manuscript where a description of the interaction is available. Some interactions are not specified (-).

at that cycle an embryonic volume 50 percent larger than the volume of the zygote cell.

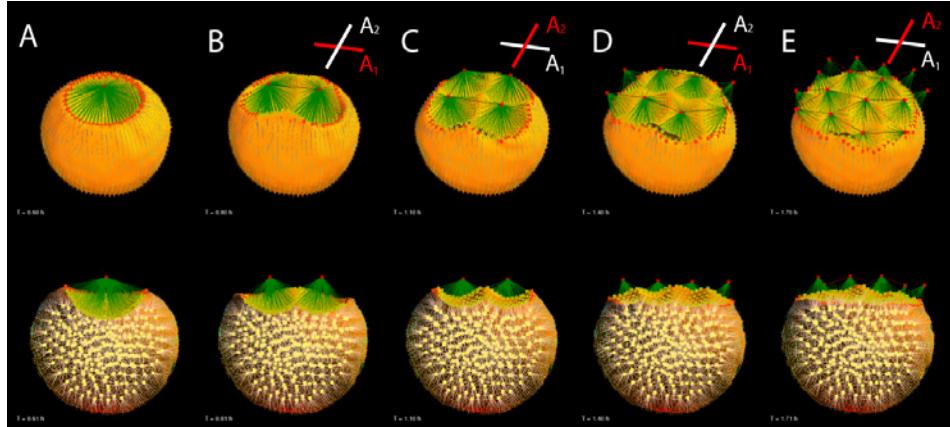


Figure 8.31: Division orientation in the model from cycle 1 to cycle 4. DC particles are represented by red cubes, YI in yellow, marginal YM in dark orange. Only the YM located above the margin are shown, the rest of the yolk membrane is represented by an orange surface. Top row: external view of the simulated embryo. The orthogonal axes A_1 and A_2 specify the axes of the division, which alternate during the first 4 cycles (red indicates the last axis). Bottom row: sagittal section of the 3D embryo rendering. A: zygote stage. B: 2-cell stage. C: 4-cell stage. D: 8-cell stage. E: 16-cell stage. A dynamical rendering is available in Movie 8.5.

The 5th division (would-be Fig. 8.31F) is oriented along the animal-vegetal axis, denoted here A_{AV} , forming two superposed layers of 16 cells. From there on, two populations co-exist in the model: the external layer of the blastoderm and the deep cell population lying between this external layer and the yolk. To avoid multiplying the terminology we shall call “EVL” the external layer, even though, strictly speaking, it has not quite differentiated into a proper EVL yet. At the 32-cell stage, the proto-EVL comprises 28 cells and the DC population comprises 4 cells. During the next 6 cell cycles, and up to cell cycle 10, each division in the EVL may contribute cells to the DC population. The successive axes of division are A_1 at cycle 6, A_2 at cycle 7, A_{AV} at cycle 8, A_1 at cycle 9, and A_{AV} at cycle 10. After that, both daughter cells of each division remain in the EVL layer. Additionally, exceptional radial divisions may also occur at cell cycle 11 (Fig. 8.32), but these will not be taken into account in the model of this case study.

At cell cycle 10, the ratio of the number of EVL cells over the number of DC particles is roughly 1/4. To reflect this, we implemented a rule guiding the contribution of dividing EVL cells to the DC population: for each cell i belonging to the EVL, if the angle formed by the dipole of daughter cells and the vector normal to the

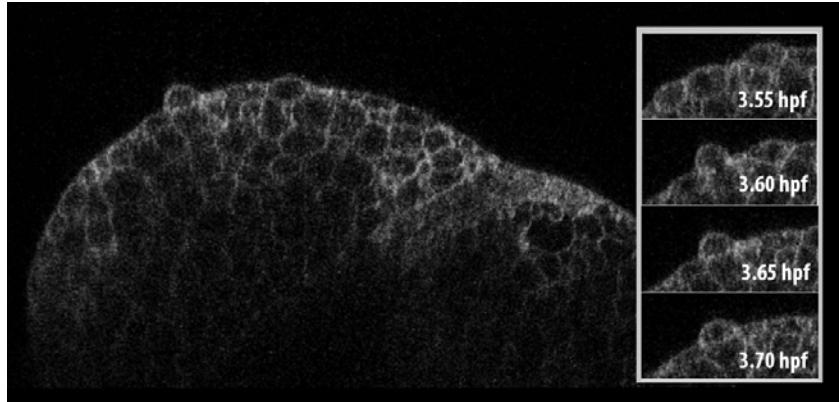


Figure 8.32: Exceptional (and last) radial cell divisions in the EVL at the cycle 11 (1024 to 2048 cells). These events are not taken into account in the present model, and shown here only as complementary information. Raw image from a live specimen imaged at the oblong stage (3.65 hpf). Insets on the right focus on one particular radial cell division, which spawns a daughter cell into the DC population and keeps the other in the EVL compartment. However, the great majority of cell divisions happening at the same time in the EVL (but not visible here) are oriented tangentially to the embryo surface.

external surface $\vec{U}_{E,i}$ (specified in the subsections below) is less than $\pi/3$, then the “inner” daughter cell is donated to the DC population. In addition, it may happen that an EVL (resp. DC) particle lies in the DC (resp. EVL) population. As this situation can create an unstable dynamics, we implemented a control device based on an “alpha shape” polyhedron. In 3D, an alpha shape is a collection of triangles covering a finite set of points, formalizing the notion of the “shape” of an ensemble of 3D points (a standard algorithm can be found in the Computational Geometry Algorithms Library, CGAL). Here, we compute the alpha shape of the cell mass composed of the union of EVL + DC + YM particles (more precisely, the YM cells that are just above the margin) once per cell cycle. If an EVL (resp. DC) particle belongs to the interior (resp. exterior) of that shape, then its type is switched to DC (resp. EVL). In the schematic visualization of Fig. 8.30A, this would correspond to a blue cross changing into a red circle or vice-versa.

EVL-EVL particle interactions

Spatial neighborhood

On its path to differentiation and becoming a proper EVL proper, the external layer (which we abusively call “EVL” here) acquires a specific behavior, distinct from that of the DC particles. EVL cells appear to deform throughout gastrulation,

stretch tangentially to the blastoderm surface and flatten in the radial direction. Thus EVL cell shapes depart more and more from the spheroidal approximation. The cell neighborhood principle introduced here will allow the model to manage such a deformation (and will be used in the next case study, Section 8.4).

The definition of the neighborhood between EVL cells is very similar to the one introduced in Section 3.2.2: first, a metric list of neighbors $\mathcal{N}_i^{m,E}$ is established; second, a topological filtering is obtained by applying the Gabriel rule (Fig. 8.33A and Movie 8.6). The main difference here is that, in order to obtain a denser neighborhood, the diameter of the Gabriel circle is shrunk by a factor $\chi^{t,E}$. Thus the modified topological neighborhood of EVL cells $\mathcal{N}_i^{t,E}$ is defined by:

$$\mathcal{N}_i^{t,E} = \left\{ j \in \mathcal{N}_i^{m,E} : \forall k \in \mathcal{N}_i^{m,E}, \left\| \vec{X}_k - \frac{1}{2}(\vec{X}_i + \vec{X}_j) \right\| \geq \frac{\chi^{t,E}}{2} \left\| \vec{X}_i - \vec{X}_j \right\| \right\}.$$

where all the cells indexed by i , j or k belong to the EVL. If $\chi^{t,E} = 1$, this neighborhood is the same as the default set \mathcal{N}_i^t defined in Section 3.2.2. In the following, the value of the topological factor $\chi^{t,E}$ will be set to 0.9.

As EVL cells start deforming during later stages, two radii are also introduced to describe the anisotropy of their shapes: a lateral radius $R_i^{\text{lat},E}$ representing the cell radius in the plane of the EVL, and an apico-basal radius $R_i^{\text{ab},E}$ representing the cell thickness. Once the topological neighborhood is established, the outward-pointing normal vector $\vec{U}_{E,i}$ of every cell i is calculated by averaging the normal vectors of all the triangles composing the topological neighborhood $\mathcal{N}_i^{t,E}$ (which generally is an irregular hexagon, sometimes pentagon, sometimes heptagon; Fig. 8.33B).

Interaction forces

We model the interaction forces among EVL cells as a pure elastic force:

$$\vec{F}_{ij}^E = -k_E(r_{ij} - r_{ij}^{\text{eq},E}) \vec{u}_{ij}, \quad \text{where} \quad r_{ij}^{\text{eq},E} = c_{2D}^{\text{eq}}(R_i^{\text{lat},E} + R_j^{\text{lat},E}),$$

k_E is the tangential stiffness constant of the EVL, and c_{2D}^{eq} is the square root of the atomic packing factor in 2D: $c_{2D}^{\text{eq}} = (\pi/(2\sqrt{3}))^{1/2} \simeq 0.9523$ (see Section 3.2.3, Fig. 3.21).

DC-EVL particle interactions

Spatial neighborhood

Neighborhoods between deep cells and EVL cells must be defined differently since, in later stages, a single EVL cell may be in contact with a dozen of underlying DC particles. First, the DC-EVL metric neighborhoods of a DC particle i and an EVL cell j , denoted respectively by $\mathcal{N}_i^{m,D:E}$ and $\mathcal{N}_j^{m,E:D}$, are obtained by visiting

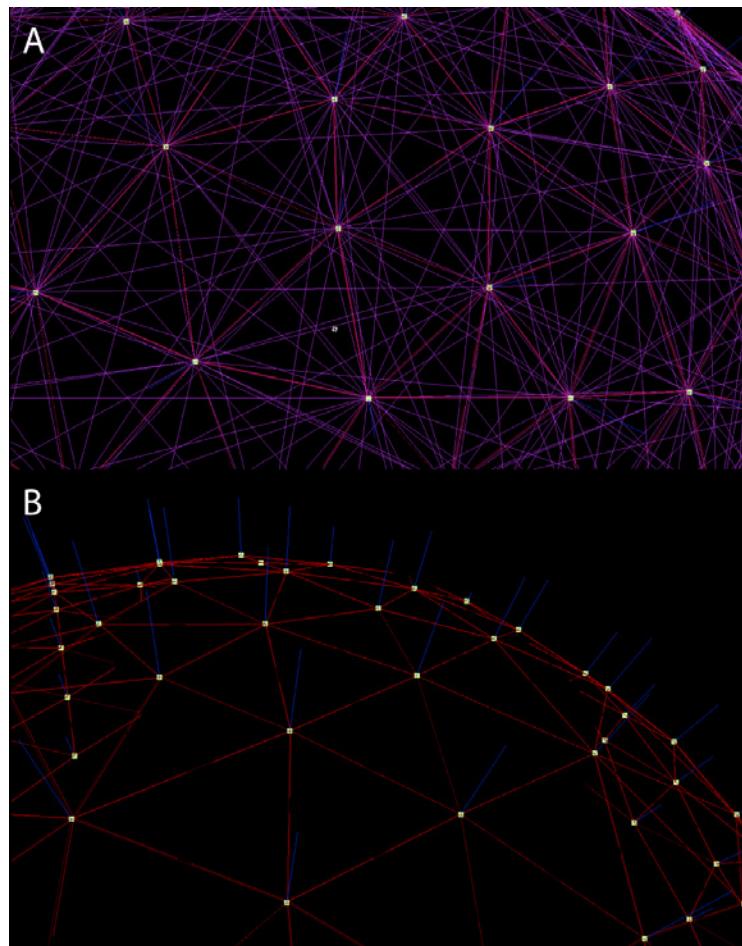


Figure 8.33: Typical EVL neighborhoods. A: Superposition of the metric neighborhood $\mathcal{N}_i^{m,E}$ (purple edges) and topological neighborhood $\mathcal{N}_i^{t,E}$ (red edges) on each EVL cell. B: The outward-pointing normal vectors $\vec{U}_{E,i}$ (blue lines) are obtained by averaging the normal vectors all the triangles composing the topological neighborhood of cell i .

each DC particle and determining its nearest EVL cell (in 3D Euclidean distance). Therefore, the D:E and E:D neighborhood sets are highly asymmetrical, since the former are singletons (of one EVL cell) while the latter may contain dozens of DC particles. Then, given these metric neighborhoods, the topological neighborhoods are obtained as follows: each DC particle i is revisited and the unique edge it has with its nearest EVL cell j is maintained if and only if there exists no DC neighbor k “above i ” in the direction perpendicular to the EVL. In other words, it means calculating the scalar product between each unitary vector \vec{u}_{ik} and the normal vector $\vec{U}_{E,j}$ and discarding edge \overline{ij} if any of these products is greater than 0.75 (Fig. 8.34):

$$\mathcal{N}_i^{t,D:E} = \left\{ j \in \mathcal{N}_i^{m,D:E} : \exists k \in \mathcal{N}_i^m, \frac{\vec{X}_k - \vec{X}_i}{\|\vec{X}_k - \vec{X}_i\|} \cdot \vec{U}_{E,j} \geq 0.75 \right\}.$$

where \mathcal{N}_i^m is the standard metric neighborhood among deep cells. A similar but asymmetrical formulation holds for each EVL cell j :

$$\mathcal{N}_j^{t,E:D} = \left\{ i \in \mathcal{N}_j^{m,E:D} : \exists k \in \mathcal{N}_i^m, \frac{\vec{X}_k - \vec{X}_i}{\|\vec{X}_k - \vec{X}_i\|} \cdot \vec{U}_{E,j} \geq 0.75 \right\}.$$

Interaction forces

Between a DC particle i and an EVL particle j , the interaction forces have the same expression as the standard “passive” attraction/repulsion forces among DC particles (see Section 3.2.3), but oriented here along the EVL normal vector $\vec{U}_{E,j}$:

$$\vec{F}_{ij}^{D:E} = A_{ij} \cdot \vec{F}_{ij}^{D:E,\text{lin}} = \begin{cases} -w_{\text{rep}}^{\text{D:E}}(r_{ij} - r_{ij}^{\text{eq}}) A_{ij} \cdot \vec{U}_{E,j} & \text{if } r_{ij} < r_{ij}^{\text{eq}} \\ -w_{\text{adh}}^{\text{D:E}}(r_{ij} - r_{ij}^{\text{eq}}) A_{ij} \cdot \vec{U}_{E,j} & \text{if } r_{ij} \geq r_{ij}^{\text{eq}} \text{ and } r_{ij} < r_{ij}^{\text{max}} \\ \vec{0} & \text{if } r_{ij} \geq r_{ij}^{\text{max}} \end{cases}$$

where r_{ij}^{eq} and r_{ij}^{max} are obtained by using the apico-basal radius $R_j^{\text{ab},E}$ instead of R_j in the case of the EVL cell:

$$\begin{aligned} r_{ij}^{\text{eq}} &= c_{\text{eq}}(R_i + R_j^{\text{ab},E}) \\ r_{ij}^{\text{max}} &= c_{\text{max}}(R_i + R_j^{\text{ab},E}) \end{aligned}$$

(with the standard coefficients c_{eq} and c_{max}), and A_{ij} is formulated as in Section 3.2.3.

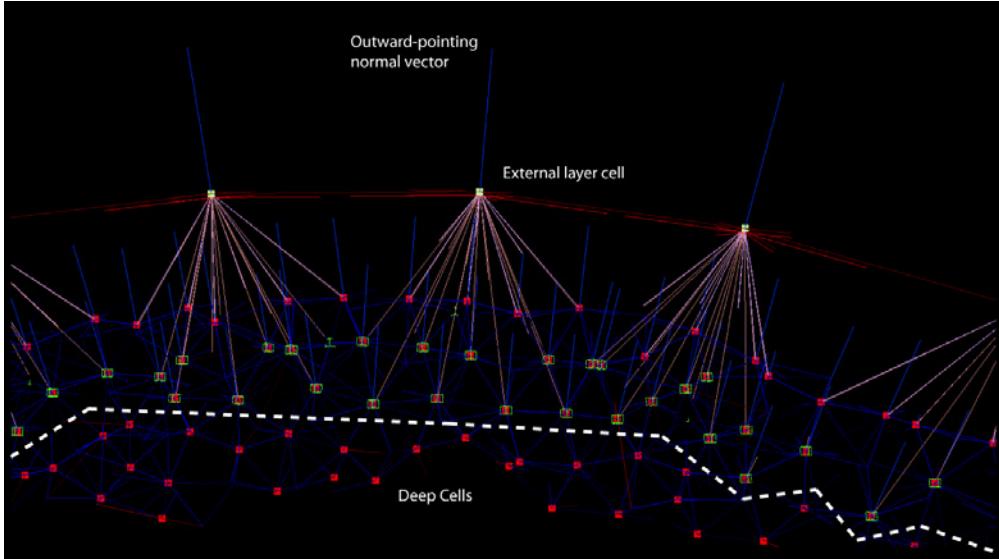


Figure 8.34: Spatial neighborhood between the EVL cells and the deep cells. This view is a 2D slice (not a 3D perspective) perpendicular to the EVL. It illustrates a late developmental stage to emphasize the size disparity between EVL cells and deep cells. Three EVL cells are displayed with their topological EVL-EVL neighborhood links (red edges). All the red cubes below are the deep cells. Topologically, we distinguish three types of cells: Below the dashed lines are the cells too distant from the EVL. Above the dashed line are the DC particles that belong to the metric neighborhoods of EVL cells $\mathcal{N}_j^{m,E:D}$. The cells surrounded by a green box are the ones that are eliminated by the topological rule: each one contains in its regular DC-DC metric neighborhood \mathcal{N}_i^m a neighbor k “higher above” in the normal direction. The top layer of deep cells, are the final topological neighbors of the EVL cells. The light purple lines represent the topological neighborhoods and the beige lines the metric neighborhoods.

EVL-YM particle interactions

Spatial neighborhood

Where the dome-shaped EVL is attached to the yolk, it forms a circular band called the *margin*. In this case study, every yolk membrane particle belonging to the margin is specially labeled and will be denoted MYM for “margin” YM (Fig. 8.35B). The spatial neighborhoods between MYM and EVL particles are the standard sets defined in Section 3.2.2 (Fig. 8.35A). The margin is expected to move “down” the surface of the yolk during epiboly and gastrulation. The mechanisms determining the conditions for margin motion will be explained in Section 8.5. They are not used in this section, which is limited to the phase that precedes epiboly.

Interaction forces

The interaction force between an EVL particle j and a YM particle i (which is by definition MYM) will be denoted by the exponent 'ym:E' and is purely elastic:

$$\vec{F}_{ij}^{\text{ym:E}} = -k_{\text{ym:E}}(r_{ij} - r_{ij}^{\text{eq,ym:E}}) \vec{u}_{ij}, \quad \text{where} \quad r_{ij}^{\text{eq,ym:E}} = 0.5(R_{\text{ym}} + R_j^{\text{ab,E}}),$$

and the stiffness coefficient is a fraction of the EVL stiffness coefficient: $k_{\text{ym:E}} = 0.25 k_E$.

DC-YM particle interactions

Spatial neighborhood

Finally, we define the links between the yolk membrane particles and the deep cells. This neighborhood has to be specified differently throughout the simulation as the DC particles are initially larger than the YM cells but eventually become smaller than them through cell divisions (typically the shift occurs at cell cycle 13). During the first period (DC larger than YM), the neighborhood is fairly simple as for every YM cell the nearest DC particle is selected. No topological filtering is added. During the second period (DC smaller than YM), the neighborhood procedure is the same as in the EVL-DC interactions. For this, we define an outward-pointing YM normal vector $\vec{U}_{\text{ym},j}$ similar to the EVL normal vector $\vec{U}_{E,j}$ and also evaluated at every simulation time step (Fig. 8.36). The topological neighborhood for the DC-YM interface is modeled after the DC-EVL interactions, too, thus for each DC particle i we have:

$$\mathcal{N}_i^{t,D:\text{ym}} = \left\{ j \in \mathcal{N}_i^{m,D:\text{ym}} : \exists k \in \mathcal{N}_i^m, \frac{\vec{X}_k - \vec{X}_i}{\|\vec{X}_k - \vec{X}_i\|} \cdot \vec{U}_{\text{ym},j} \leq -0.75 \right\}.$$

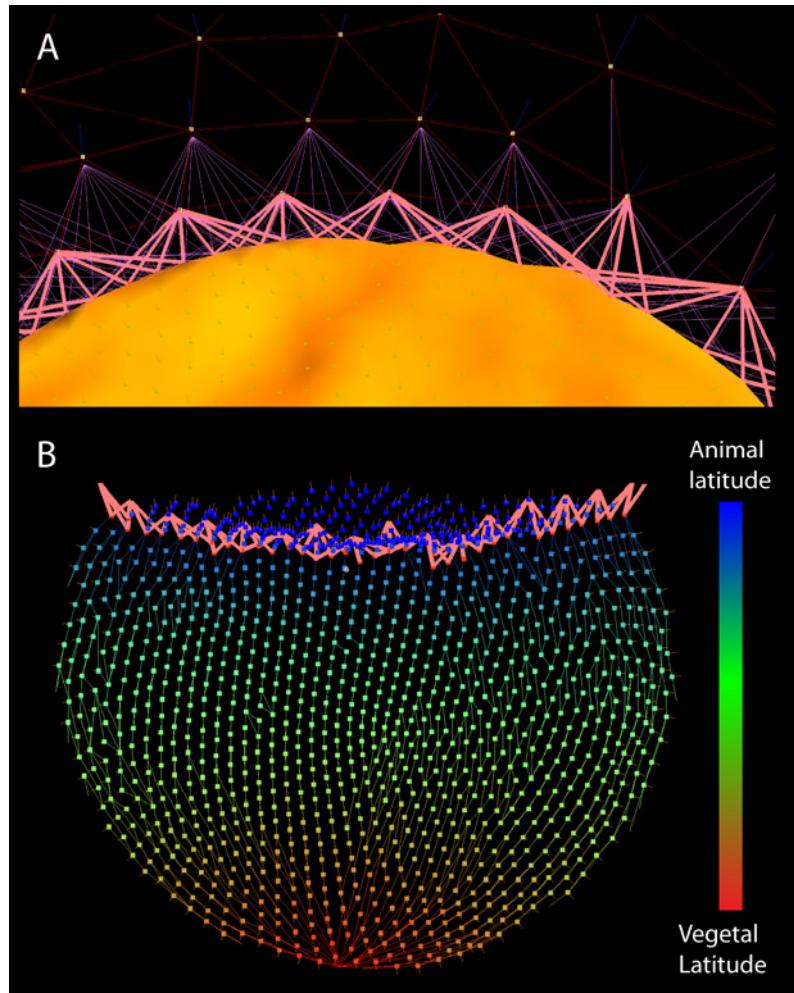


Figure 8.35: Region of attachment of the EVL to the yolk, or “margin”. A: Metric (thin purple lines) and topological (fat pink lines) neighborhoods between EVL cells and the MYM particles. B: Concentric yolk marginal circles. From the vegetal pole (bottom) to the animal pole (top) of the yolk, YM particles are classified into 25 concentric latitudes, which are displayed here with a color map from red to blue. The pink lines on top are the same as in A, but the EVL cells are hidden.

and for each YM particle j :

$$\mathcal{N}_j^{t,\text{ym:D}} = \left\{ i \in \mathcal{N}_j^{m,\text{ym:D}} : \exists k \in \mathcal{N}_i^m, \frac{\vec{X}_k - \vec{X}_i}{\|\vec{X}_k - \vec{X}_i\|} \cdot \vec{U}_{\text{ym},j} \leq -0.75 \right\}.$$

Note that the sign and the limit of the scalar product is reversed as the YM normal vector $\vec{U}_{\text{ym},j}$ is now pointing toward the cells. To simplify, we will denote by $\mathcal{N}_i^{\text{D:ym}}$ and $\mathcal{N}_j^{\text{ym:D}}$ these neighborhoods, whether they are of the metric type (first phase) or topological type (second phase).

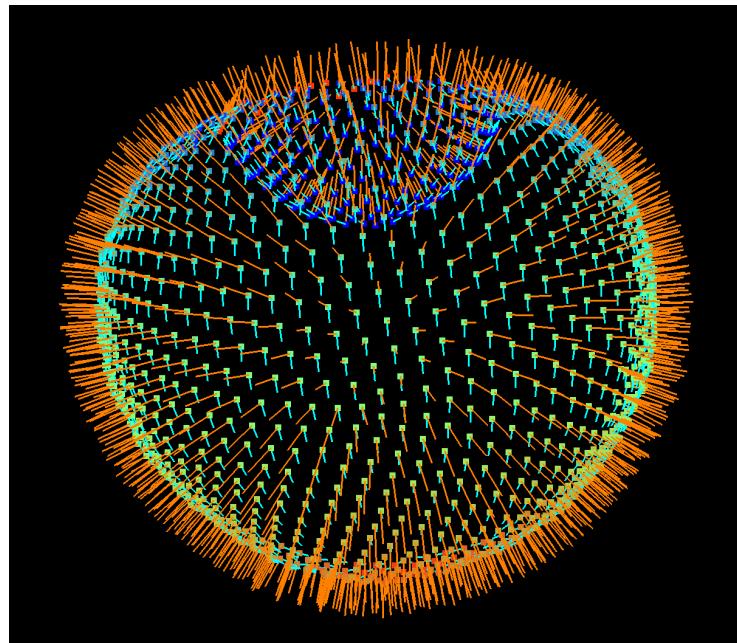


Figure 8.36: Yolk membrane particles' normal and tangential vectors. At each time step, each YM particle j updates its outward-pointing normal vector $\vec{U}_{\text{ym},j}$ (orange lines). It also updates a tangential vegetal-pole-oriented vector denoted by $\vec{U}_{\text{ym},j}^{\parallel}$ (blue lines), which will be used later on, in the fifth case study (Section 8.5).

Interaction forces

The interaction forces exerted between YM and DC particles are of the standard “passive” attraction-repulsion type, formulated in a way similar to DC-EVL, but using the unitary edge vectors \vec{u}_{ij} :

$$\vec{F}_{ij}^{\text{D:ym}} = A_{ij} \cdot \vec{F}_{ij}^{\text{D:ym,lin}} = \begin{cases} -w_{\text{rep}}^{\text{D:ym}}(r_{ij} - r_{ij}^{\text{eq}})A_{ij} \cdot \vec{u}_{ij} & \text{if } r_{ij} < r_{ij}^{\text{eq}} \\ -w_{\text{adh}}^{\text{D:ym}}(r_{ij} - r_{ij}^{\text{eq}})A_{ij} \cdot \vec{u}_{ij} & \text{if } r_{ij} \geq r_{ij}^{\text{eq}} \text{ and } r_{ij} < r_{ij}^{\text{max}} \\ \vec{0} & \text{if } r_{ij} \geq r_{ij}^{\text{max}} \end{cases}$$

where r_{ij}^{eq} and r_{ij}^{max} are obtained by using the constant radius R_{ym} (defined in Section 8.1.1) instead of R_j in the case of the YM cell:

$$\begin{aligned} r_{ij}^{\text{eq}} &= c_{\text{eq}}(R_i + R_{\text{ym}}) \\ r_{ij}^{\text{max}} &= c_{\text{max}}(R_i + R_{\text{ym}}) \end{aligned}$$

(with the standard coefficients c_{eq} and c_{max}), and A_{ij} is formulated as in Section 3.2.3.

Parameter space for the biomechanical properties of the blastula

We hypothesize that from the 8-cell stage until the high stage, the cell-cell adhesion coefficient w_{adh} , the cell-yolk adhesion coefficient $w_{\text{adh}}^{\text{ym:D}}$ and the EVL stiffness coefficient k_E are the major parameters underlying blastoderm shaping within a single dynamical regime. By exploring the parameter space defined by these three parameters, we expect to find the appropriate conditions for a realistic simulation of this phenomenon that matches the quantitative data obtained from live imaging. We examine this question in the next section.

8.3.2 Simulation, Parameter Space and Validation

Prior to a more systematic exploration of the 3D parameter space defined by the triplet $(w_{\text{adh}}, w_{\text{adh}}^{\text{ym:D}}, k_E)$, we carried out a first confrontation of the simulated and reconstructed data based on a first set of parameters that was hand-picked to produce a blastoderm width and height roughly fitting the observations shown in the introduction of Section 8.3, Fig. 8.29. We implemented a simultaneous visualization of simulated data and live data from [254] (Movie 8.7). Three data sets from that paper, with IDs 081014h, 080917h and 081024h, correspond to live imaging of zebrafish embryos throughout the first 10 cell divisions with THG signal as shown in Fig. 8.2 and SHG signal revealing mitotic spindles. In addition to the reconstructed cell lineage for each of these data sets, the paper proposes a prototypic embryo calculated from 6 different specimens by averaging cell dipole lengths at the time of cell division, and cell division orientations.

We designed two measures to support the fitness function:

- Measure 1 is based on the calculation of 3D alpha shapes. At each time step, we compute the 3D alpha shape corresponding to the set of embryo cell centers.

- Measure 2 is based on the calculation of 2D alpha shapes. These are achieved by a rotational projection of the particle set around the animal-vegetal (AV) axis, each cell being assigned a position on the half-plane (Fig. 8.37). This operation provided measurements easier to analyze than the 3D alpha shape.

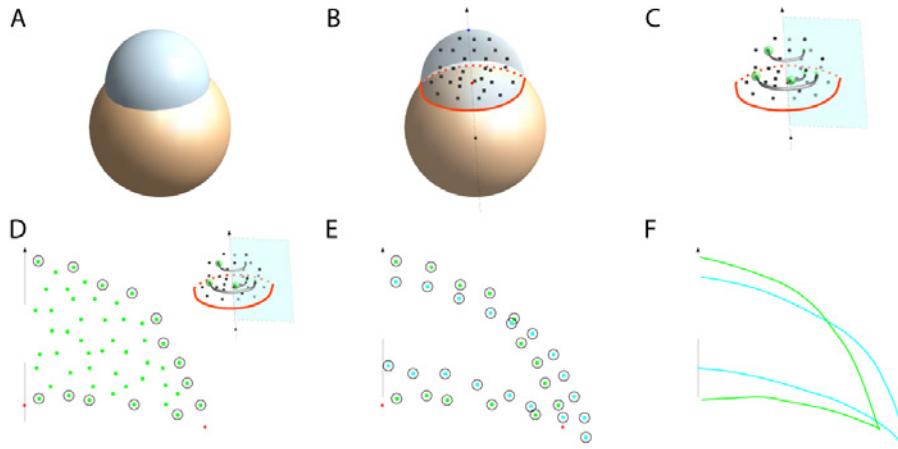


Figure 8.37: A 2D alpha shape strategy for assessing the embryo shape evolution. A: schematic view of the embryo, with the blastoderm area in blue and the yolk in beige. B: The vertical arrow indicates the animal-vegetal (AV) axis, the red line delineates the blastoderm margin, i.e. the interface between the yolk, the blastoderm and the environment. C: Each cell position is projected on a half-plane by rotation around the AV axis. D: A 2D alpha shape yields the positions of the cells at the blastoderm surface, characterizing the contour of the embryo. E: Two different 2D shapes from two embryos are superimposed. F: The discrete shapes are interpolated via analytical functions.

A possible measurement to characterize the embryo shape and then compare different embryos is to infer the blastoderm contours, in order to both obtain its outer surface and its inner surface in contact with the yolk. Each curve can be described with a limited number of geometric parameters, and we can then predict the shape changes through the evolution of these parameters. This strategy is relevant to characterize both the evolution of the embryo's shape and the shape differences between two embryos at the same stage.

Fitness Function

We now propose one fitness function for each type of measure:

- Fitness 1 is calculated by counting the number of cells from the simulated data set that lie outside the alpha shape of the experimental data set, and vice versa,

then adding these two numbers. The best scenario naturally corresponds to a zero sum, i.e. no outliers on either side.

- Fitness 2 is calculated by the adding the differences between the geometric parameters describing the outer and inner fitting curves of the blastoderm shape.

The proposed measures and fitness are being currently implemented and tested. We do not have results to show yet.

8.3.3 Discussion

Alternatively to the two measures proposed, one could also think of a third one: We should be able to obtain from the BioEmergences reconstruction workflow the algorithmic segmentation of the outer and inner surfaces of the blastoderm from 3D+time imaging of live embryos, either by THG or by 2PEF of fluorescently stained cell membranes. This strategy should provide the best approximation (validated by eye inspection) of the blastoderm shape.

When it is finished, we expect from this study an optimal shape of the blastoderm that will be used as an initial condition for assessing the properties underlying further development (i.e. doming and epiboly).

We also expect some indication about the homogeneity of the process. In other words, we should be able to propose whether the blastoderm shape evolution is adequately described by a single set of parameters, or whether it goes through different transient phases, possibly indicating changes at a more microscopic level in terms of cell adhesion, cell tension, cell intrinsic motility, cell yolk interaction, and external layer differentiation into EVL.

Further exploration of the blastula biomechanics will be proposed through the study of cadherin mutants or morphants. During the blastula period, the *cdh1* morphant, with impaired cadherin-1 expression, both at the cell-cell interface and the cell-yolk interface, have an early lethal phenotype. This phenotype is typically an early blastula shaping defect and we briefly discuss here how our model would allow us to account for the disruption of the blastoderm as early as the 4-cell stage under perturbed cell-cell and cell-yolk adhesion (Fig. 8.38). We hypothesize that such a phenotype cannot be obtained if cell-cell adhesion is homogeneously decreased in the mutant. We rather propose to account for the mutant phenotype to implement an heterogeneous field of adhesion coefficients w_{adh} and $w_{\text{adh}}^{\text{ym:D}}$ and we expect to observe similarities with the live phenotypes described by Babb et al. [313].

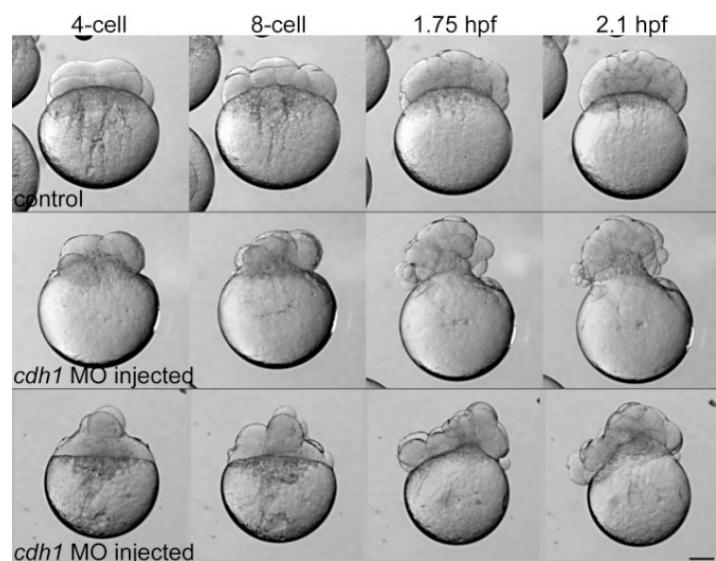


Figure 8.38: Early cleavage patterns and compaction are disrupted in injected embryos. Zebrafish embryos were injected with control morpholino oligonucleotide (MO), or *cdh1* MO, and monitored by time-lapse microscopy. In control embryos (top panel), cleavage planes were regular, blastomeres were approximately equal in size, and cells were strongly adherent and formed a compact mass atop the yolk cell. In *cdh1* MO-injected embryos (bottom two panels), the orientation of cleavage planes between blastomeres was irregular, cell size was unequal, adhesion between blastomeres was reduced, and the embryos did not compact normally. hpf: hours post-fertilization. Scale bar 100 μ m. Image and caption from [313]

8.4 Cell Behaviors in the Enveloping Cell Layer Compartment

The enveloping cell layer (EVL), which differentiates during the blastula stages, has long been hypothesized to be an extra embryonic compartment. Recent studies rather suggest a striking parallel between the EVL and mammal epidermis [314]. The EVL reinforces its epithelial properties through differentiation steps. EVL biomechanical properties are expected to be essential for further gastrulation steps, during which the EVL is hypothesized to stand increasing tension, undergo limited positional rearrangements [265], and increase the apical and basal surfaces greatly (see Fig. 8.39 for an illustration of EVL cells' increase in *Fundulus*). Various models and measurement methods of the epithelium have been elaborated during the past years (see Section 2.1.4). However, most studies concerned active monolayered epithelia, supposedly performing intrinsic active cell behaviors. In contrast, the zebrafish EVL is supposed to passively react to external stress applied at the level of the margin. In this case study, we explore the specificities of the zebrafish EVL in terms of cell rearrangement and propose a mechanical principle for the regulation of cell proliferation.

8.4.1 Hypotheses and Model

Our implementation of the EVL formation is detailed in the previous case study (see Section 8.3.1). In our model, all the blastoderm cells divide synchronously for the first 10 cell divisions. From the 11th cell cycle, while the deep cells' cycle slows down and desynchronizes in a geometrical fashion (see second case study, Section 8.2), EVL cells have a radically different behavior. We hypothesize that the cell cycle is not determined by an internal clock anymore, but rather by the action of an external *tangential tension force* exerted on the EVL cells. The division is controlled by the size of the cells, which is itself a function of the mechanical forces exerted on them. In this study, the overall external forces exerted on the EVL are stretching forces applied at the level of the margin, where the EVL cells are attached to the yolk. We assume that these stretching forces might be triggered by the action of the deep cells. The latter mechanism is studied in the next case study (Section 8.5), which focuses on the same time period, i.e. epiboly phase 1. Here, the external stretch is considered given and will not be detailed.

External pressure

We model the external tangential tension exerted over an EVL cell i as a “pressure” P_i^E , which will be generally negative to express a tension but can also be positive in the case of a compression. It is calculated by adding the amplitudes of the forces exerted on the various “facets” of i by its neighboring EVL cells j and dividing by

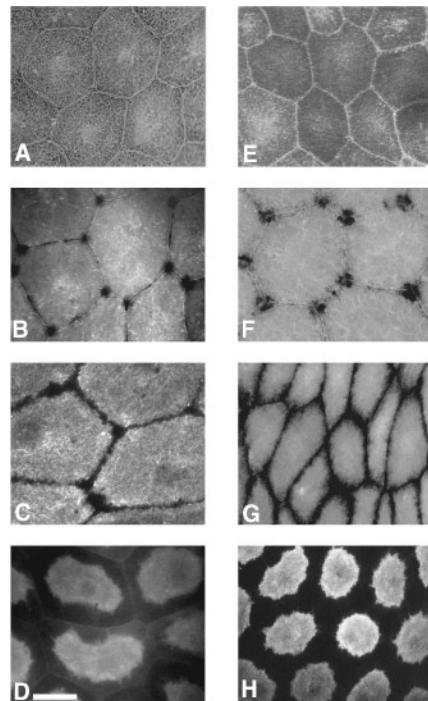


Figure 8.39: Regional membrane turnover within the apical membranes of *Fundulus* enveloping layer cells. Pairs of photographs showing the temporal progression of membrane turnover (seen as loss of fluorescence) of EVL cells labeled with fluorescent lectin (A-D) and lipid (E-H). Embryos stained with fluorescent lectin initially show a uniform pattern of fluorescence (A). Six hours later, apical membrane turnover has removed the fluorescent markers from regions where three or more cells come in contact. This loss of fluorescence continues, and by 18 hr (C) after staining, lectin turnover has spread from the cellular contact points along the cell margins. By 36 hr after lectin staining, fluorescence is restricted to a central domain (D). This same pattern of membrane turnover is seen when embryos are labeled with fluorescent lipid: (E) uniform staining at Time = 0 hr; (F) loss of fluorescence at cellular contact points 12 hr after staining; (G) formation of a nonfluorescent peripheral domain 29 hr after staining; and (H) lipid turnover 42 hr after staining creates a peripheral band that extends 10-12 μm inward from cell boundaries. Scale bar, 30 μm . Image and caption from Fink and Cooper, 1996 [264]

the number of facets:

$$P_i^E = -\frac{1}{n_i} \sum_{j \in \mathcal{N}_i^{t,E}} (\vec{F}_{ij}^E \cdot \vec{u}_{ij})$$

where $n_i = \text{card}(\mathcal{N}_i^{t,E})$ is the number of EVL neighbors of cell i . We recall that the EVL-EVL interaction force \vec{F}_{ij}^E was modeled as a pure elastic force in the previous case study (Section 8.3.1):

$$\vec{F}_{ij}^E = -k_E(r_{ij} - r_{ij}^{\text{eq},E}) \vec{u}_{ij}, \quad \text{where } r_{ij}^{\text{eq},E} = c_{2D}^{\text{eq}}(R_i^{\text{lat},E} + R_j^{\text{lat},E}),$$

k_E is the tangential stiffness constant of the EVL, and c_{2D}^{eq} is the square root of the atomic packing factor in 2D: $c_{2D}^{\text{eq}} = (\pi/(2\sqrt{3}))^{1/2} \simeq 0.9523$ (see also Section 3.2.3, Fig. 3.21). Therefore, putting everything together, the pressure reads:

$$P_i^E = \frac{k_E}{n_i} \sum_{j \in \mathcal{N}_i^{t,E}} (r_{ij} - c_{2D}^{\text{eq}}(R_i^{\text{lat},E} + R_j^{\text{lat},E}))$$

EVL cell growth

EVL cells are modeled as deformed cylindrical-shaped particles, i.e. they are described by a lateral radius $R_i^{\text{lat},E}$ and an apicobasal “half-height” $R_i^{\text{ab},E}$, also called radius for simplicity here. Since we hypothesize that the volume of an EVL cell is conserved between two divisions, these two radii are related by $R_i^{\text{ab},E}(R_i^{\text{lat},E})^2 = \text{cst}$. Thus we will employ the expression “cellular growth” in the EVL to refer to an augmentation of the lateral radius only—concurrently with a corresponding decrease of the apicobasal radius.

The growth rate of the EVL cells’ lateral radius $R_i^{\text{lat},E}$ is expressed as follows:

$$\frac{dR_i^{\text{lat},E}}{dt} = \begin{cases} -\gamma_E R_i^{\text{lat},E} & \text{if } P_i^E \geq \theta_E^+ \text{ (positive pressure: compression)} \\ 0 & \text{if } \theta_E^- P_i^E \theta_E^+ \\ +\gamma_E R_i^{\text{lat},E} & \text{if } P_i^E \leq \theta_E^- \text{ (negative pressure: tension)} \end{cases}$$

where θ_E^- (resp. θ_E^+) is a negative (resp. positive) pressure threshold, and γ_E is the growth ratio. In this study, the EVL cells are never compressed by their neighborhood, so the influence of the positive pressure will not be explored. The value of its threshold is set to 1.

Divisions of the EVL cells

In this model, EVL cells divide after the 11th cell cycle if their lateral radius exceeds a limit radius denoted by $R_{\text{lim}}^{\text{lat},E}$. We calculate this limit radius via a coefficient $c_{\text{lim}}^{\text{lat},E}$ linking it to the lateral radius of the EVL cells after cycle 11, denoted by $R_{g11}^{\text{lat},E}$:

$$R_{\text{lim}}^{\text{lat,E}} = c_{\text{lim}}^{\text{lat,E}} R_{\text{g11}}^{\text{lat,E}}.$$

The orientation of the axis of cell division is picked randomly in the EVL tangential plane. The division is considered symmetrical, i.e. the two daughter cells have an equal lateral radius, and since their surface area is roughly equal to half of their mother's, we get: $R_{\text{daughter}}^{\text{E}} = R_{\text{mother}}^{\text{E}}/\sqrt{2}$.

The parameter space for the EVL cells behavior

In summary, the parameter space is 3-dimensional and comprises:

- the limit radius coefficient $c_{\text{lim}}^{\text{lat,E}}$
- the EVL cellular growth ratio γ_{E}
- the “negative pressure” threshold θ_{E}^-

8.4.2 Simulation, Parameter Space and Validation

Characterizing the epithelial structure has become a popular topic in the biological systems modeling community (see Section 3.2.4). To our knowledge, however, this type of study has not been realized yet for the zebrafish EVL, most certainly due to a lack of quantitative data. Thanks to the work of our team on the BioEmergences platform (described in Section 7.1), we now possess a quantitative temporal evolution of cell neighborhoods in the animal pole region of the EVL during phase 1 epiboly (Fig. 8.40).

Reconstruction of EVL neighborhoods from live imaging

The live specimen used in this study is a wild-type embryo injected with *mCherry/H2B* and *egfp-ras* RNAs and imaged with a Leica SP5 upright microscope MLSM (see Section 7.1.1). First, the cell lineage was automatically reconstructed through the BioEmergences workflow (see Section 7.1.2). The processed data goes from 3h48min to 5h43min (hpf), i.e. from the onset of epiboly to the germ-ring stage. Then, for each time step t , a subpopulation \mathcal{E}_t of the EVL cells was corrected manually, both in the positions of the nuclei's approximate centers and in the cell tracking (yellow cubes in Fig. 8.40A and Movie 8.8). 26 mitoses occurred in the selected cell population.

In addition to the cell lineage, a reconstruction of the cells' neighborhood edges was also performed. A first automated step was realized via the Delaunay triangulation algorithm (pink lines in Fig. 8.40A). Then, the neighborhood edges between EVL cells belonging to the subpopulation \mathcal{E}_t of the swarm were manually corrected and validated with the Mov-It software (yellow/green area). The task was facilitated by highlighting the membrane signal with a simple technique: a convex hull

of the embryo shape was calculated at each of the 100 reconstructed time steps so that the volume rendering of the membrane channel raw data was hidden inside the hull, leaving a clear view of the external part of the signal (the convex hull is the blue surface, the yellow signal gives the position and shape of the EVL cells' membrane). In the following, we denote by $\mathcal{N}^{l,E}(t)$ the corrected neighborhood of the live specimen linking the centers of population \mathcal{E}_t at the imaging times step $t = 0 \dots 100$.

Population-wide measures

The measures performed on the live specimen belongs to two categories, as presented in Escudero et al. [149]:

- *geometrical measures* based on the apical cell surface $S_i^l(t)$ of EVL cell i at imaging time step t (Fig. 8.43).
- *topological measures* based on the number of neighborhood edges $n_i^l(t)$, also called *degree*, of EVL cell i at time step t (Fig. 8.44).

The shape of the apical cell surface is approximated by a disc whose radius is the average half-length of the neighborhood edges $l_{ij}(t)$ over neighbors j in $\mathcal{N}_i^{l,E}(t)$, thus:

$$S_i^l(t) = \pi \left(\frac{1}{2} \sum_{j \in \mathcal{N}_i^{l,E}(t)} l_{ij}(t) \right)^2$$

One way to characterize the temporal evolution of the EVL is to define the *average apical cell surface* at time t over the subpopulation \mathcal{E}_t of manually corrected cells (Fig. 8.43A):

$$S^l(t) = \frac{1}{\text{card}(\mathcal{E}_t)} \sum_{i \in \mathcal{E}_t} S_i^l(t)$$

Similarly to the apical surface measure, another characterization is the *average neighborhood degree* at time t (Fig. 8.44B):

$$n^l(t) = \frac{1}{\text{card}(\mathcal{E}_t)} \sum_{i \in \mathcal{E}_t} n_i^l(t)$$

A third classical topological measure is *frequency distribution* D_k^l , $k \in [3..9]$ of the neighborhood degrees (i.e. the number of cells that have a given degree k) over the population \mathcal{E} and the period of 100 time steps. It is plotted as a histogram in Fig. 8.42.

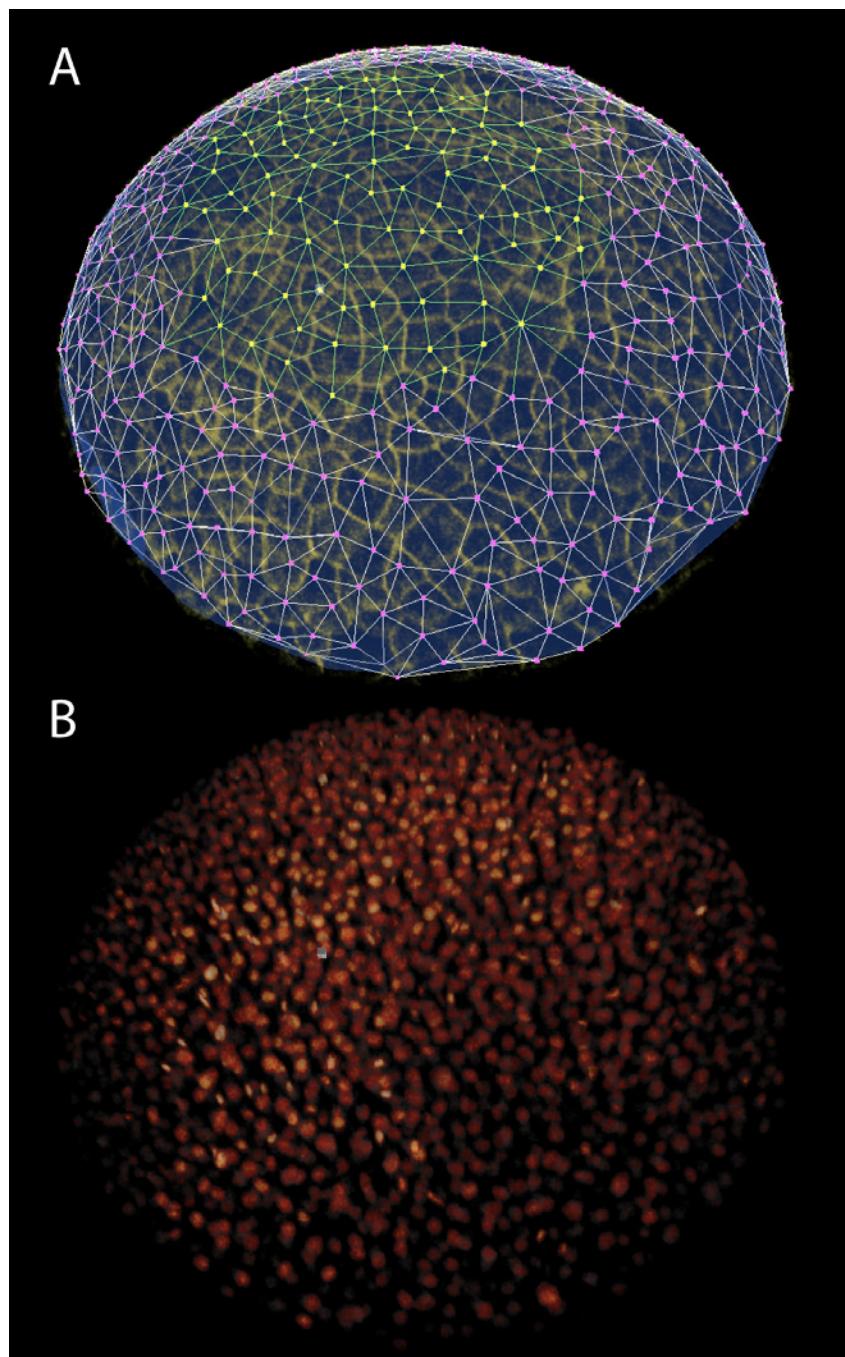


Figure 8.40: Reconstruction of the EVL from 3D+time imaging. Both images belong to the same dataset 070418a and have been generated by the Mov-It software. A. Superposition of the convex hull (blue surface) calculated from the swarm of cell centers, the volume rendering of the cell membrane channel (yellow signal), the cell positions (cubes) and the cell neighborhood edges (lines connecting the cubes). The manual validation and correction of the cell positions, cell tracking and cell neighborhoods of the yellow/green population were corrected by Adeline Boyreau, the pink population corresponds to the automated reconstruction. B: Volume rendering of the nucleus channel.

Mitosis-related measures

In addition to the population-wide measures, we introduce local measures of the average apical surface $S_{\text{mit}}^l(t')$ and neighborhood degree $n_{\text{mit}}^l(t')$ before and after each of the 26 *mitoses* occurring in the manually corrected tissue. The timeline is rescaled so that the mitoses are occurring at $t' = 0$, and the rescaled time values t' are comprised between -50 and 50 . We denote by \mathcal{M}^l the set of mitoses in the live specimen. Each mitosis $m \in \mathcal{M}^l$ is described by a pair of values: $m = (t_m, i_m)$, where t_m is the time step of the mitosis, and i_m the id of the mitotic cell. As a mitosis produces two daughter cells, each mitosis is counted twice, once for each daughter id (i.e. the cardinal of the set \mathcal{M}^l is 52). The measure of the neighborhood degree $n_{\text{mit}}^l(t')$ around a mitotic cell i is obtained by aligning the cells' degrees $n_i^l(t)$ at the time of each mitosis (Fig. 8.41) and dividing them by the number $\#$ of cells for which there exists a measure at t' (Fig. 8.43B and 8.44B):

$$n_{\text{mit}}^l(t') = \frac{1}{\#(t')} \sum_{m \in \mathcal{M}^l} n_{i_m}^l(t' + t_m)$$

where $\#(t')$ is the cardinal of the set of measures (surface or degree) that are defined at the time step t' . For example, Fig. 8.41B, $\#(0) = 10$, $\#(-20) = 8$, and $\#(50) = 6$.

The average surface around mitoses $S_{\text{mit}}^l(t')$ is expressed in a similar fashion, but integrates another factor in its expression in order to deal with the heterogeneity of cell surfaces observed in the data. We divide the surface of a mitotic cell by its new surface just after mitosis:

$$S_{\text{mit}}^l(t') = \frac{1}{\#(t')} \sum_{m \in \mathcal{M}^l} \frac{S_{i_m}^l(t' + t_m)}{S_{i_m}^l(1 + t_m)}$$

We expect in further work to extend this type of study to a small cohort of specimens. The one image here with the shortest possible time step (67s) until the 6-somite stage showed photobleaching of the *mCherry* by the end of gastrulation and some cell death in the brain by 24 hpf. Although we only use here the first hours of the time lapse, we cannot exclude that the features of this specimen are far from being prototypical. In any case, we observe that the variability of the EVL cells' surface is high (large error bars in Fig. 8.43A). This is somehow surprising if we assume that the cell cycles remained synchronous for the first 10 divisions (see second case study, Section 8.2) and given that the thickness of the EVL cells seems to be homogeneous at a given time step. In this context, and if the cell volume does not increase, the data would be best explained by asymmetric divisions. The increase in EVL cell surface is expected and accelerates during epiboly phase 1 (Fig. 8.43A).

The measure of the cells' average apical surface around mitosis $S_{\text{mit}}^l(t')$ shows a higher slope before division than after it (Fig. 8.43B). This could indicate that

smaller cells are less sensitive to the overall pressure in the tissue, and thus do not expand their surface as fast as the larger ones. We also observe that, surprisingly, the average surface of the daughter cells just after division is more than half the surface of the mother cell (approx. 60 percent). It could be due to a bias in the spheroid approximation of the measure, or to a thinning of the cell as early as the time of cytokinesis.

The average evolution of the degree (i.e. number of neighbors of a given cell) through time remains constant (Fig. 8.44A). However, we observe that the average degree of the dividing cells diminishes after mitosis, which is expected, but does not recover its pre-mitosis value during the hour following the division (Fig. 8.44B). Such a behavior indicates that the tissue undergoes little rearrangement during the phase 1 epiboly. We compared the characteristics of the zebrafish EVL to other animal epithelia from the literature (Fig. 8.42B). The histogram of the distribution of degree for the zebrafish EVL is radically different from the narrower histogram provided by so-called proliferating epithelia. In Gibson et al. [144], the *Drosophila* wing disc, the *Hydra* epidermis and the *Xenopus* tail epidermis are shown to have similar distribution of degrees. This type of distribution is typical of epithelia that are shaped by proliferation only. In the chick embryo primitive streak, where cells are intrinsically motile, the degree distribution is much broader and favors lower numbers of neighbors (see chick streak in Fig. 8.42B) [148]. The zebrafish EVL distribution seems closer to the chick streak case.

Simulated specimens

The simulated embryos are modeled in a simulated time corresponding to the period measured in the live embryos: from 3h52min to 5h43min simulated time. The driving mechanism of the epiboly is the protrusive behavior of the deep cells of the blastoderm. This mechanism is explored in detail in the fifth case study, Section 8.5. The measures performed in the live embryo are reproduced for each simulated embryo. Simulated embryos are denoted by the s superscript instead of the l superscript. Whereas the simulation time step is shorter than the live specimen imaging time step (6 simulated seconds per time step versus 67 seconds), the measures were realized with the same time interval as the measures in the live specimen.

Fitness Function

Similarly to the other case studies, we define one objective subfunction per measure mentioned above, in order to compare the live and simulated embryos:

- F_S , the average surface evolution function:

$$F_S = \sum_{t=1}^{100} \|S^l(t) - S^s(t)\|$$

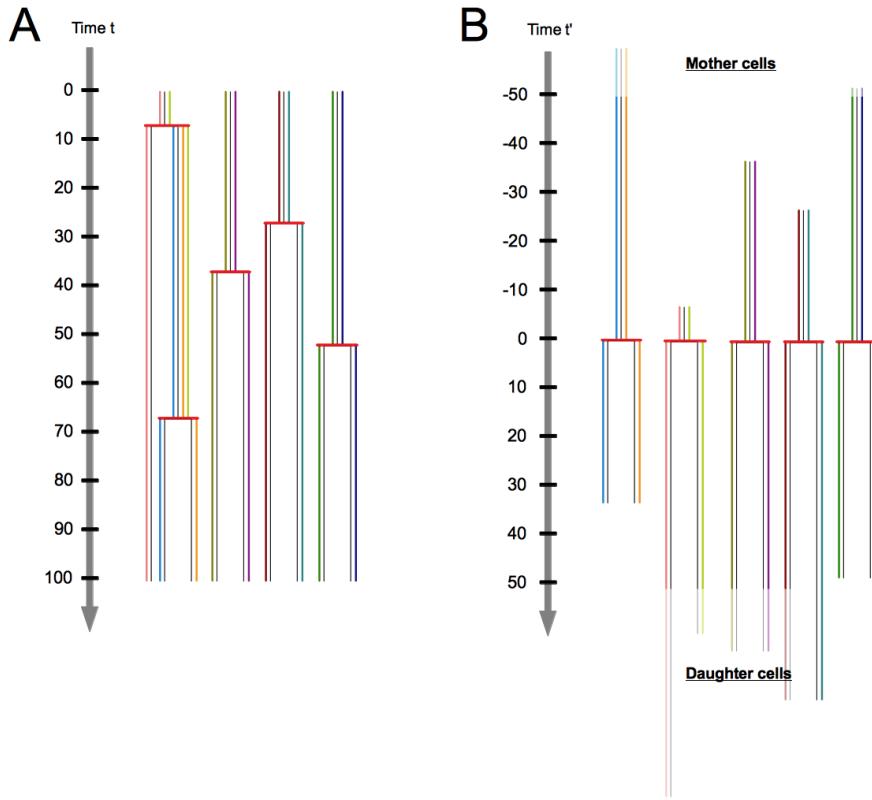


Figure 8.41: Temporal alignment of cell divisions. The abstract examples of mitosis displayed to illustrate the strategy do not belong to the real measured population \mathcal{M}^l . The colors identify individual mitoses m of \mathcal{M}^l in both panels. A: 5 mitoses occurring at various time steps. B: The same mitoses translated in the new timeline t' . The measured quantities represented by the colored lines are available for each time step t' . The quantities are considered if they belong to the $[-50, 50]$ interval. Beyond this interval, the discarded values are covered by a white overlay.

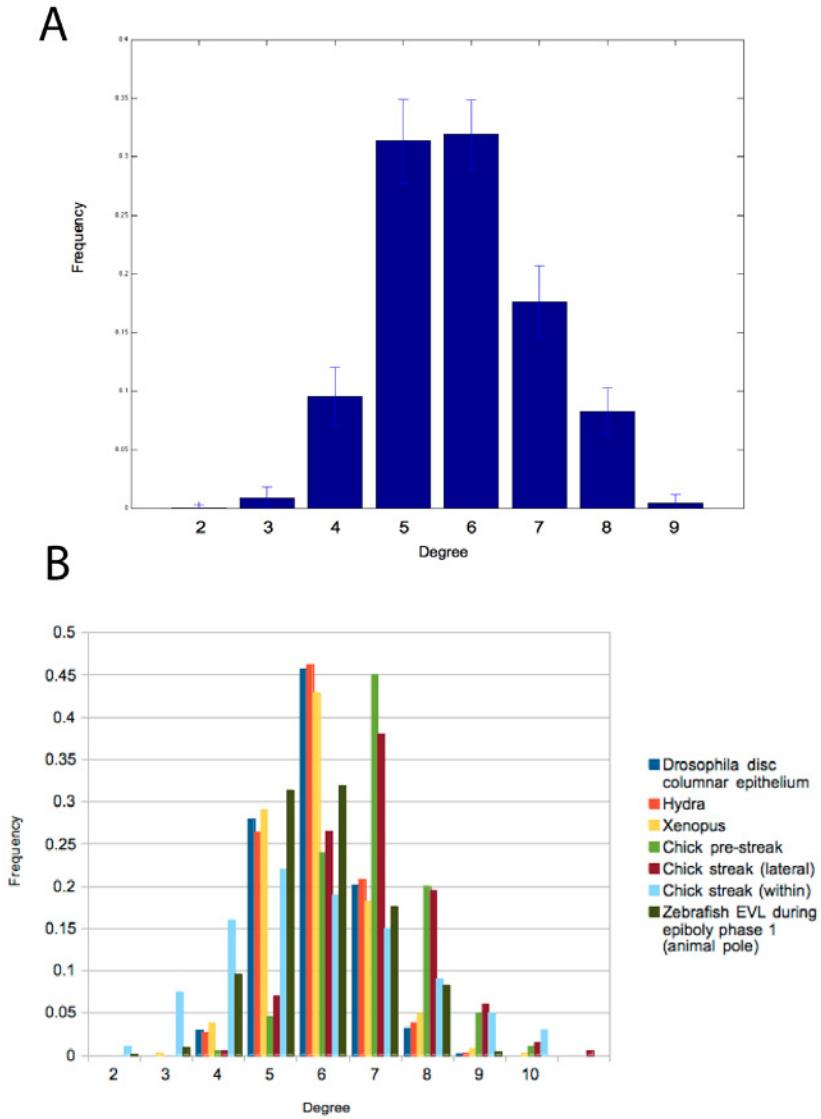


Figure 8.42: Histogram of the number of neighbors in the whole population \mathcal{E} and comparison with other tissues. A: In ordinate, the frequencies of the degree D_k^l observed in the corrected population \mathcal{E} . Error bars indicate the standard deviation for each degree's frequency. Measures calculated and plotted by Paul Villoutreix. B: Comparison of the distributions of degrees in the *Drosophila* wing disc (blue, 2,172 cells), the *Hydra* epidermis (orange, 602 cells), the *Xenopus* tail epidermis (yellow, 1,051 cells), the chick pre-streak (green, 351 cells), the chick lateral streak (brown, 634 cells), the chick streak (pale blue, 613 cells), and the zebrafish enveloping layer during epiboly phase 1 (dark green, 5725). Data for *Drosophila*, *Hydra* and *Xenopus* are extracted from Gibson et al. [144], data for the chick epithelia are extracted from Sandersius et al. [148].

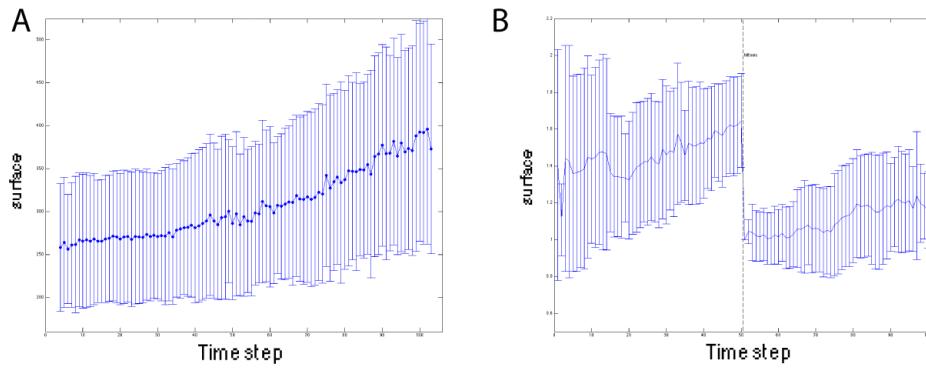


Figure 8.43: Evolution of the EVL cells' average surface.

A: Temporal evolution of the average cell surface $S^l(t)$ in the population \mathcal{E} . The measure starts at time step 4 (3h52min) and ends at time step 103 (5h43min). Consecutive time steps are separated by 1min07s. An increase of 50 percent is measured during this period. B: Temporal evolution of the average cell surface $S^l_{\text{mit}}(t')$ around mitosis. All 52 mitoses are superposed as if they were occurring simultaneously at time step $t' = 0$. The cells' apical surfaces are normalized so that the surface of the cell measured immediately after mitosis is equal to 1. The average and standard deviation at each time step before and after mitosis are calculated and plotted. Both measures were calculated and plotted by Paul Villoutreix.

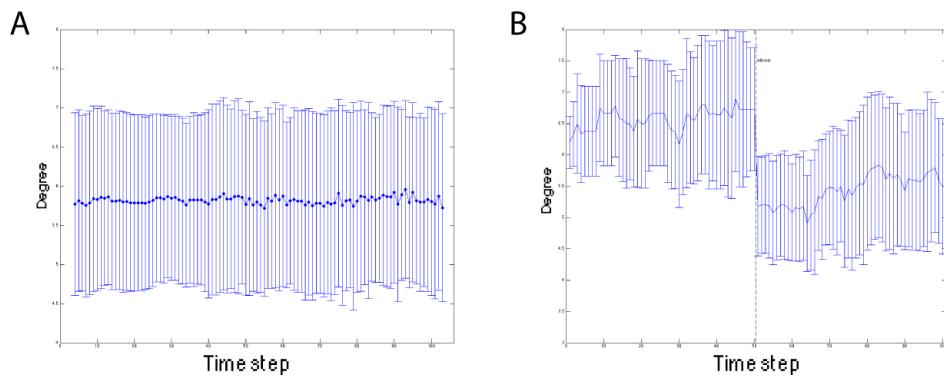


Figure 8.44: Evolution of the EVL cells' average degree. A: Temporal evolution of the averaged cell degree $n^l(t)$ in the population \mathcal{E} . The measure starts at the time step 4 (3h52min) and ends at the time step 103 (5h43min). Consecutive time steps are separated by 1min07s. The average degree remains roughly constant during this period. B: Temporal evolution of the average cell degree $n^l_{\text{mit}}(t')$ around mitosis. All 52 mitoses are superposed as if they were occurring simultaneously at time step $t' = 0$. The average and standard deviation at each time step before and after mitosis are calculated and plotted. Both measures were calculated and plotted by Paul Villoutreix.

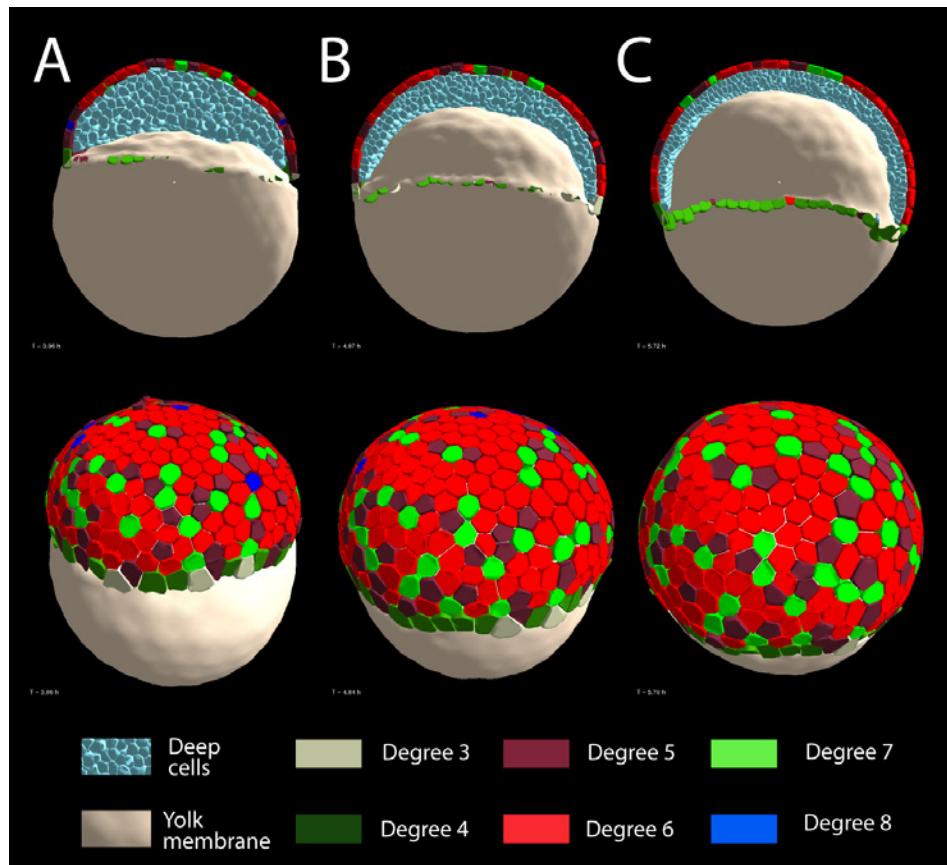


Figure 8.45: The simulated embryo's EVL during epiboly phase 1. All deep cells are colored in blue. The color of the EVL cells indicates the degree of the cell (see legend). First row: the embryo is cut along the bilateral symmetry plane, the yolk is empty as no yolk interior particles are represented. Second row: external view of the embryo. A: onset of epiboly (approx. 3.9h simulated time). B: 30-percent epiboly stage (approx. 4.85h simulated time). C: 50-percent epiboly, the internalization is not modeled here (approx. 5.7h simulated time). The movie from which these images are extracted is Movie 8.9.

- F_n , the average degree evolution function:

$$F_n = \sum_{t=1}^{100} \|n^l(t) - n^s(t)\|$$

- F_D , the average degree distribution function:

$$F_D = \sum_{k=3}^9 \|D_k^l - D_k^s\|$$

- $F_{S,\text{mit}}$, the average mitotic surface function:

$$F_{S,\text{mit}} = \sum_{t'=-50}^{50} \|S_{\text{mit}}^l(t') - S_{\text{mit}}^s(t')\|$$

- $F_{n,\text{mit}}$, the average mitotic degree function:

$$F_{n,\text{mit}} = \sum_{t'=-50}^{50} \|n_{\text{mit}}^l(t') - n_{\text{mit}}^s(t')\|$$

These five objective subfunctions are merged into a global fitness function F through a weighted sum method. As in the second case study, Section 8.2, we renormalize each objective function by its maximum value:

$$F = \frac{F_S}{\max F_S} + \frac{F_n}{\max F_n} + \frac{F_D}{\max F_D} + \frac{F_{S,\text{mit}}}{\max F_{S,\text{mit}}} + \frac{F_{n,\text{mit}}}{\max F_{n,\text{mit}}}$$

Fitness Landscape

We explored the 3D parameter space specified in 8.4.1. Parameters include the limit radius coefficient $c_{\text{lim}}^{\text{lat,E}}$, the EVL growth ratio γ_E , and the pressure threshold θ_E^- . The parameter space was regularly sampled over the following range with the following cardinalities (Fig. 8.46):

	Min.	Max.	Cardinality	Unit
$c_{\text{limit}}^{\text{lat,evl}}$	1.1	3	15	-
γ_{evl}	1.001	1.1	15	-
θ_{evl}^-	20	50	15	N

Figure 8.46: Range, cardinalities and units of the three parameters explored in this study.

A general trend of this exploration of the parameter space is the limited influence of the EVL growth ratio γ_E parameter. All the fitness landscapes show little variation along the vertical axis (Figs. 8.47-8.52). In reaction to a given external pressure, a faster expansion of the EVL cell surface will lower the overall EVL pressure faster, making it pass below the pressure threshold θ_E^- . This threshold is the key parameter controlling the resistance of the simulated embryo to external pressure. As expected, the limit radius coefficient $c_{\lim}^{\text{lat},E}$ influences the average surface evolution function F_S (Fig. 8.47). If the limit radius coefficient value is too small, the EVL cells divide a lot, leading to numerical instabilities when the pressure threshold is high (red “potholes” in Figs. 8.48-8.52). If the limit radius coefficient value is too high, the EVL cells become larger than in the live specimen, and above a value of about 2.5, the EVL cells will not divide during the epiboly phase 1 period (Fig. 8.47). The size of the EVL cells is also controlled by the pressure threshold which limits the expansion the EVL cells. The best value for the limit EVL cell radius before a division is about 1.45 times the radius at the previous cell generation (in this case the 11th one).

The topological measures highlight a limited performance of the model in terms of the characterization of the number of neighboring cells. The average degree is not sensitive to variations in parameter space (Figs. 8.48, 8.49). A closer examination of the simulated distribution of degrees shows a stable narrow distribution similar to the one observed by Gibson et al. in proliferating epithelia (histograms from simulations not shown). This low sensitivity is due to the strong constraint imposed by the neighborhood algorithm (see Section 8.3) and the spheroidal particle approximation.

The global fitness landscape offers a more contrasted evaluation of the parameter space with a domain of best fitness delimited by a value of $c_{\lim}^{\text{lat},E}$ around 1.4 and a value of θ_E^- around 40 (Fig. 8.52).

8.4.3 Discussion

The model of the enveloping layer compartment presented here offers the possibility to test the mechanical properties of the zebrafish EVL. To our knowledge, it had never been studied based on quantitative measurements as presented above. Compared with the literature, our real measures (Fig. 8.42B) tend to show that the EVL epithelium obeys a behavioral law different from the ones described so far for this kind of tissue, with a distribution of the number of neighbors situated in an intermediate position between the proliferating epithelia and the motile epithelia. The evaluation of the model provided mild results: the simulation performed well on the geometrical side, but poorly on the topological side. We hypothesize that to improve these results, the abstract level of the cell shape characterization, i.e. a flat spheroid described by a single particle, may require a refinement of the neighboring rules to account for the topological properties of the EVL. However, the overall me-

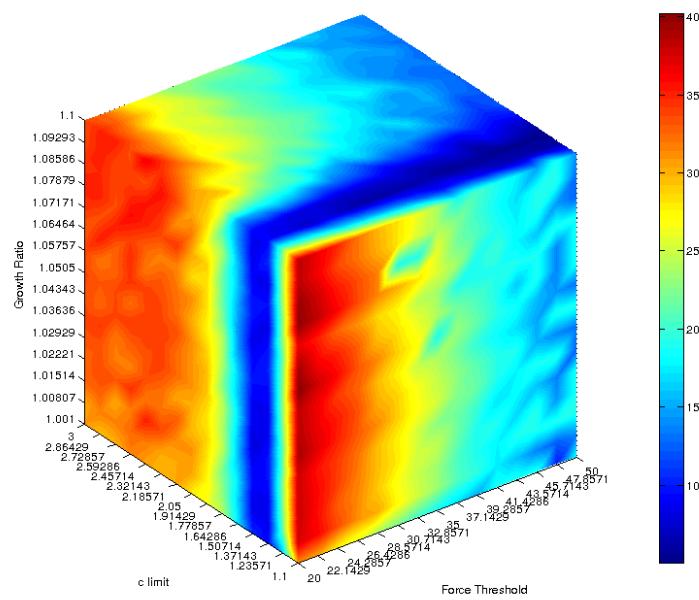


Figure 8.47: 3D plot of the average surface evolution function F_S . The vertical axis indicates the EVL growth ratio γ_E , the depth axis indicates the limit radius coefficient $c_{\lim}^{\text{lat},E}$ and the horizontal axis indicates the pressure threshold θ_E^- .

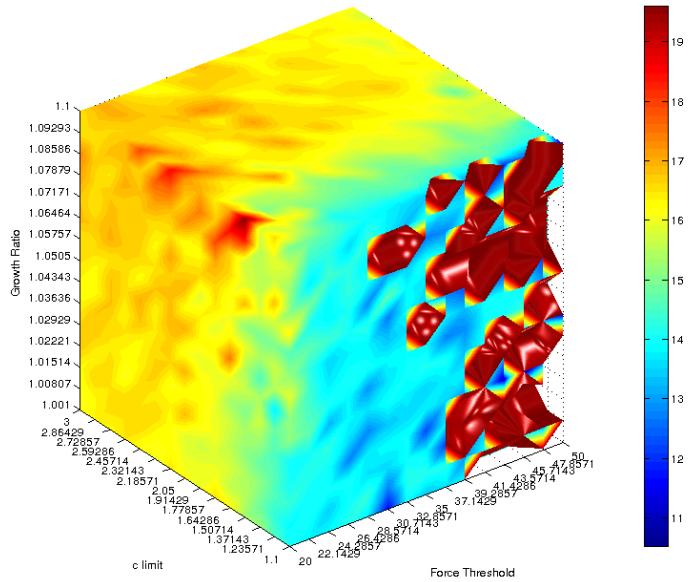


Figure 8.48: 3D plot of the average degree evolution function F_n . The red “potholes” visible on the lower limit radius coefficient $c_{\lim}^{\text{lat},E}$ face of the cube are due to simulation divergence. See Fig. 8.47 for details.

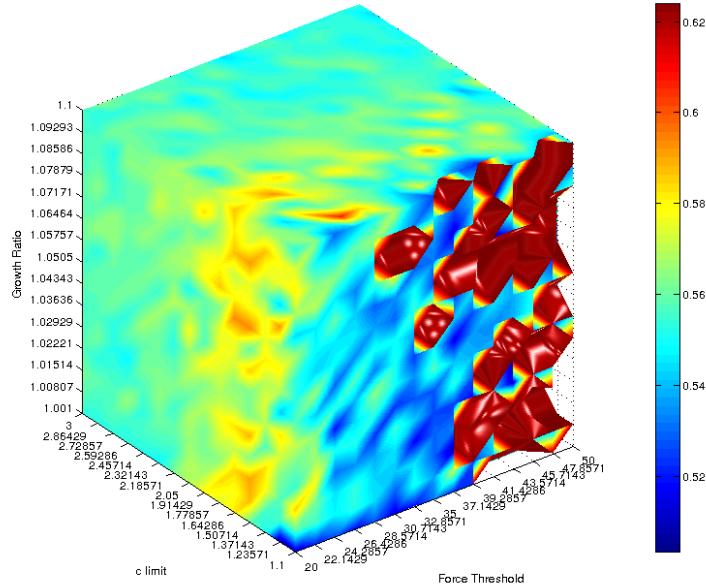


Figure 8.49: 3D plot of the average degree distribution function F_D . See Fig. 8.47 for details.

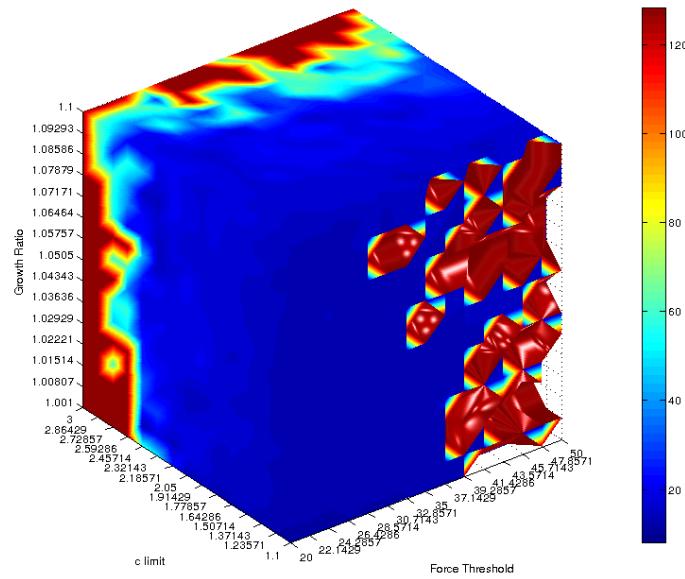


Figure 8.50: 3D plot of the average mitotic surface function $F_{S,\text{mit}}$. See Fig. 8.47 for details.

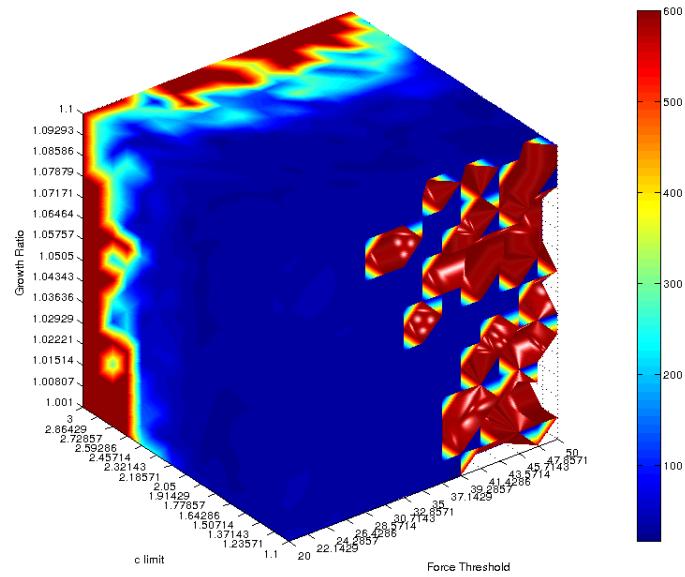


Figure 8.51: 3D plot of the average mitotic degree function $F_{n,\text{mit}}$. See Fig. 8.47 for details.

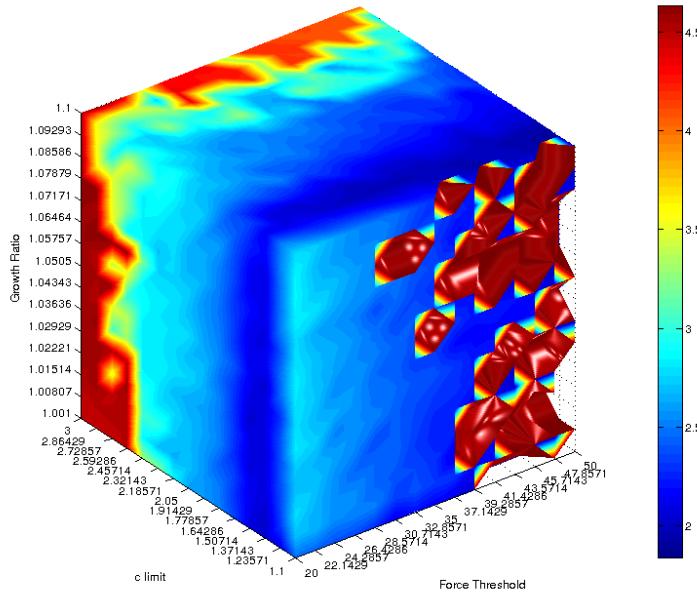


Figure 8.52: 3D plot of the global fitness function F . See Fig. 8.47 for details.

chanical properties of the tissue can indeed lead to the epibolic motion, as detailed in the next case study. The mechanism by which the margin constraint induces the pressure in the EVL is also presented in the next section.

As already mentioned, this study should be extended to a small cohort of individuals. Further investigation should also apply the strategy to the EVL at later stages, until the closure of the tail bud, and if possible on the entire EVL. We could compartmentalize the EVL in different subregions to test whether the mechanical constraints produce different behaviors according to the spatial position. A possible compartmentalization could be an animal-vegetal slicing of the EVL in concentric rings (see Movie S.10).

8.5 Intercalation Patterns

Epiboly is characterized by a flattening of the deep cell mass and its spreading over the yolk cell toward the vegetal pole (Fig. 8.53). Two successive episodes may be distinguished, as described in Chapter 6: a “first epibolic” phase from the high stage to the shield stage, with the intermediate oblong and sphere stages, and a “second epibolic” phase starting at the time of the internalization of the hypoblast. The second phase will be studied in the last case study, about “Gastrulation” (Section 8.6). We focus here on the first phase of epiboly and the cell movements involved in it. From the onset of epiboly, deep cells start to *intercalate radially*, cells from the

depth of the blastoderm moving toward the blastoderm surface. Although intrinsic deep cell behaviors are supposed to be responsible for the deformation of the embryo at these stages, we cannot exclude an active participation of the yolk syncytial layer (YSL) or the newly differentiated EVL to the epiboly movement either. In this case study, we use macroscopic measures characterizing the doming phenomenon and microscopic measures characterizing the intercalation patterns, to show that the intrinsic behavior of deep cells is sufficient to trigger upward yolk bulging (doming motion) and downward margin progression toward the yolk equatorial latitude. In this context, we will also evaluate how the YSN margin and the enveloping layer (EVL) tangential stiffness modulate the deep cells' driving force.

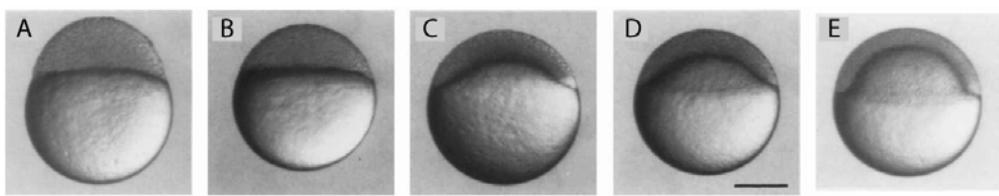


Figure 8.53: Early gastrula period. Face views. A: transition between the high and oblong stages (3.5 h). B: transition between the oblong and sphere stages (3.8 h). C: Dome stage (4.3 h). D: 30-percent epiboly stage (4.7 h). E: 50-percent epiboly stage (5.25 h). Scale bar: 250 μm . Image and caption adapted from Kimmel et al. (1995) [253]

8.5.1 Hypotheses and Model

The initial state of the simulation in this case study is the same as in the previous one (Section 8.4.1). The simulation starts at a simulation time equivalent to 3h30 hpf (just after the high stage [253]) and stops at a simulation time equivalent to 5h42 hpf (germ ring stage [253]). The deep cells lie on top of the yolk, stuck between the YSL and the EVL (Fig. 8.55A).

Passive margin sliding

We envision five possible, but non-exclusive scenarii explaining how the margin moves toward the vegetal pole: (i) the internal YSL (iYSL, see Fig. 6.2) actively spreads over the yolk (in its cortical region), carrying both the deep cells and the EVL margin with it, (ii) the EVL actively spreads over the blastoderm, carrying the margin with it, (iii) the margin is pulled downward by an active mechanism in the external YSL (eYSL), (iv) some active mechanism inside the yolk triggers the convex bulge inside the blastoderm area, and (v) the deep cells actively intercalate and their collective behavior induces a pressure at the marginal region, pushing the resisting margin toward the vegetal pole.

Scenarii (i) and (iii) would require some active mechanism in the microtubules' network linking the yolk syncytial nuclei (YSN), either by pushing from the iYSL or pulling from the vegetal eYSL. Disruption of the microtubules with nocodazole at the sphere stage is not sufficient, however, to stop the epibolic motion [262]. Scenario (ii) would require an active flattening of the EVL apico-basal thickness and conjugated extension of its lateral surface. This would imply that at the margin, the EVL would move toward the yolk equatorial latitude ahead of the deep cells, but this is not what we observe in our imaging data (Fig. 8.54 and Movie 8.10). The inside of the yolk has also not been described to contain a well structured cytoskeleton. This penalizes the possibility envisioned by scenario (iv) of an active mechanism occurring in this domain.

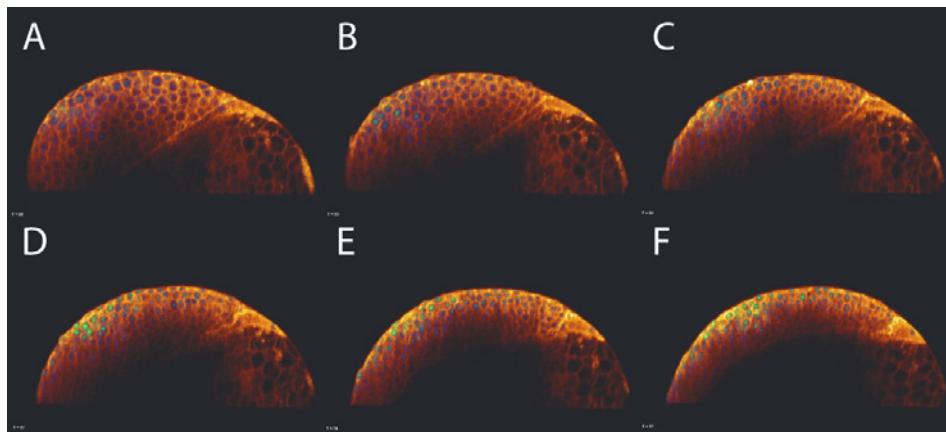


Figure 8.54: Wild type zebrafish embryo imaged live from the oblong stage to 30-percent epiboly. We observe here a slice passing approximately through the center and the animal pole of the embryo. A: oblong stage. B: transition between oblong and sphere stage. C: sphere stage. D: transition between sphere and dome stage. E: dome stage. F: transition from dome to 30-percent epiboly stage. Dataset ID 071222bF. The movie from which these images are extracted is available in 8.10

All the experiments described in the literature on pre-internalization embryos indicate that phase 1 epiboly is a robust process, with probably redundant underlying mechanisms. However, the only scenario which does not lead to contrasted arguments is scenario (v), thus our goal will be to show in the simulation context that the active intercalation of the deep cells is sufficient to drive the epiboly during phase 1. This scenario requires *a mechanism that will convert the push exerted by the deep cells over the margin into a sliding movement of the margin toward the vegetal pole*. We describe here how this mechanism is modeled.

Force model

The marginal deep cells are stuck in the corner formed by the YSL and the EVL. We expect that the margin will slide toward the vegetal pole if the norm of the tangential force exerted by the deep cells (DC) on the marginal yolk membrane particles (MYM), denoted by $\vec{F}_{ij}^{\text{m:D},\parallel}$, is larger than a given “resistance threshold” $\theta_{\text{m},\parallel}$ (Fig. 8.55C). We call this force the “pushing force” as it expresses the localized quantity of force exerted by a DC particle j over an MYM particle i . Its equation reads:

$$\vec{F}_{ij}^{\text{m:D},\parallel} = \begin{cases} -w_{\text{rep}}^{\text{ym:D}}(r_{ij}^{\parallel} - r_{ij}^{\text{eq},\parallel}) A_{ij}(r_{ij}^{\parallel}, R_i, R_{\text{ym}}) \cdot \vec{U}_{\text{ym},i}^{\parallel} & \text{if } r_{ij}^{\parallel} < r_{ij}^{\text{eq},\parallel} \\ \vec{0} & \text{if } r_{ij}^{\parallel} \geq r_{ij}^{\text{eq},\parallel} \end{cases}$$

where the repulsion coefficient $w_{\text{rep}}^{\text{ym:D}}$ is the same as the one controlling the repulsion at the YM-DC interface (see Section 8.3.1), A_{ij} is the surface of contact described in Section 3.2.2, r_{ij}^{\parallel} is the scalar product of the relative position vector $r_{ij}\vec{u}_{ij}$ and the tangential vector $-\vec{U}_{\text{ym},i}^{\parallel}$ of the MYM particle i (Fig. 8.55C,D), and $r_{ij}^{\text{eq},\parallel} = c_{\text{eq}}(R_i + R_{\text{ym}})$. Note that only the repulsive part of the force (rr^{eq}) has an non-zero formulation because we do not consider here the reverse situation of marginal deep cells going back toward the animal pole of the yolk and pulling the margin with them. Naturally, that case could also be implemented if needed (for example to account for a mutant phenotype).

Chemical model

The objective of this study is to show that radial intercalation is sufficient to drive the epiboly phase 1. At the cellular level, the mechanism attributed to the cells is the *bipolar protrusion*. We propose two similar means of specifying radial polarization fields through the Waddingtonian Timeline Specification described in Section 4.2:

- The first polarization field is specified by the ligand-based polarization mode ‘(a)’ based on a diffusive ligand Q_1 secreted by the EVL. For this ligand, the YSL acts as a sink (Fig. 8.56, top embryo).
- The second polarization field is obtained by reversing the sink and source roles base on another ligand Q_2 (Fig. 8.56, bottom embryo).

If the EVL-deep cells and the YSL-deep cells interfaces were exactly parallel during phase 1 epiboly, these two gradients would generate equal polarization fields. This is obviously not the case before 30 percent epiboly (see Movie 8.10 before $T = 90$), justifying our choice of two different radial polarization fields. It should be noted that using a ligand-based polarization mode should not be interpreted as any commitment to how a polarization field is generated in the real embryo. For this

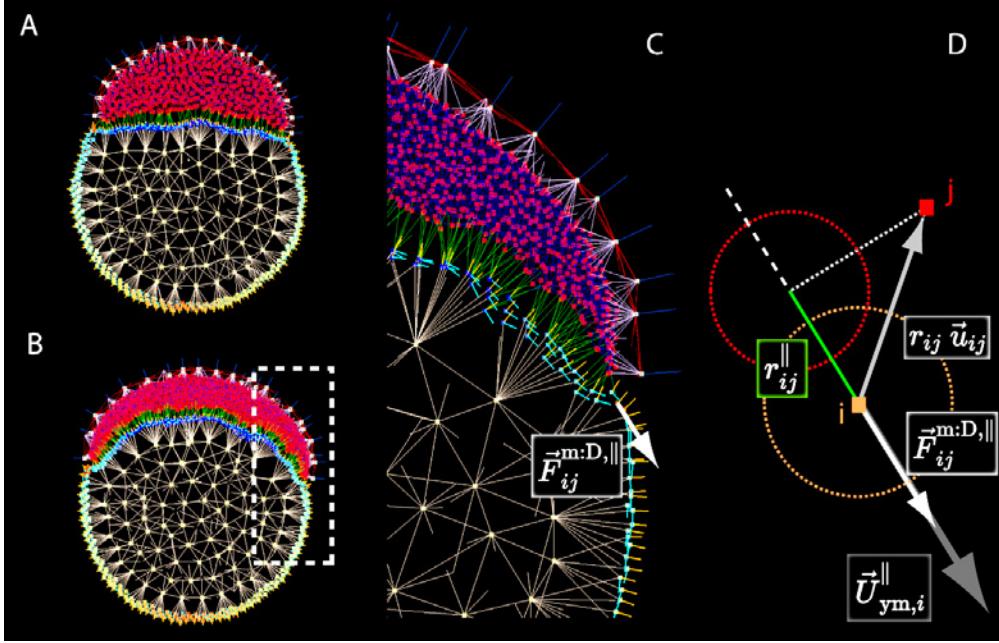


Figure 8.55: Mechanism controlling the passive sliding of the blastoderm and EVL margin toward the vegetal pole.

A: Sagittal section of the simulated embryo at the onset of epiboly.

B: Sagittal section at 30 percent epiboly. C: Zoom on the marginal region of the embryo at 30 percent epiboly. When a deep cell (DC, red) j is in contact (green lines) with a margin yolk membrane (MYM, orange) particle i , a “pushing force” $\vec{F}_{ij}^{m:D,\parallel}$ is calculated to estimate the mechanical pressure exerted by j on the margin at the cellular level. If this force exceeds a certain “resistance threshold” $\theta_{m,\parallel}$, the MYM particle loses its marginal properties, and transmits it to one or several regular YM particle(s) at a more vegetal latitude. C: Schema describing the pushing force exerted by DC particle j over MYM cell i . The pushing force is non-zero only if the distance between the positions of j and i projected on the tangential vector $\vec{U}_{ym,i}^{\parallel}$ (not shown) is smaller than the equilibrium distance $r_{ij}^{eq,\parallel} = c_{eq}(R_i + R_{ym})$. The orange and red circles highlight the radii of cells j and i respectively.

study, we remain “agnostic” with respect to the actual physico-chemical polarization mechanisms, and are interested only in the effects that such abstract fields have on cell movements and emerging morphogenetic processes at a more macroscopic level.

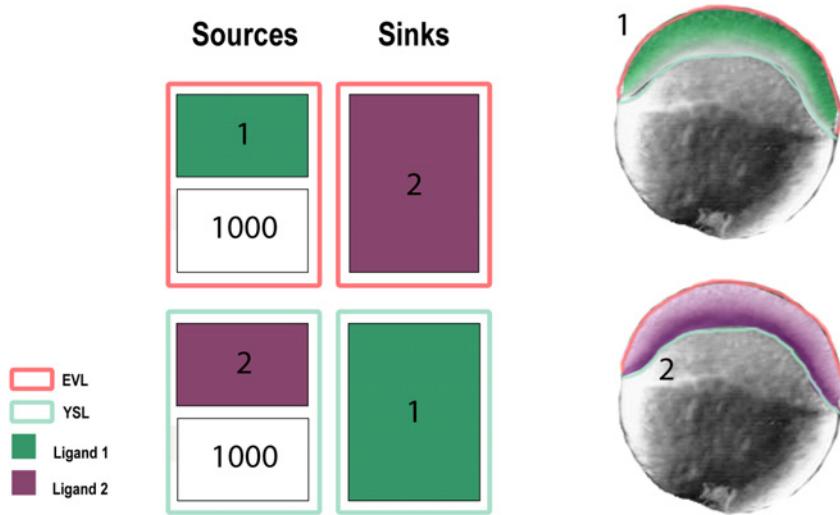


Figure 8.56: Ligand diffusion table L for the “Intercalation” case study. Two gradient fields are established by ligand diffusion across the dome region formed by the deep cells. They contribute to the specification of the axes of the polarization: a ligand Q_1 is secreted by the EVL with secretion rate $s_1 = 1000$ and absorbed by the yolk membrane particles, more precisely the YSL. Conversely, a ligand Q_2 is secreted by the YSL with rate $s_2 = 1000$ and absorbed by the EVL.

The polarization fields specified above are very regular, or certainly less fluctuating than would the case biologically. To test the effect of some stochasticity on these fields, we introduced a new binary parameter λ_r controlling the randomness of the polarization field: if $\lambda_r = 0$, the polarization axes \vec{U}_i are as specified above; otherwise, if $\lambda_r = 1$, the polarization field is purely random, i.e. we add an *ad hoc* rule generating, for each cell i , a random vector $\vec{U}_{r,i}$ every 15 minutes of the simulation time. For each cell i , the equation of the effective polarization axis \vec{U}_i^e integrating this stochastic factor is:

$$\vec{U}_i^e = \lambda_r \vec{U}_r + (1 - \lambda_r) \vec{U}_i$$

followed by a renormalization step to keep the polarization vector unitary.

Force-chemical coupling model

Once the polarization field is established, the protrusive behavior of the cells needs to be parametrized. For the sake of simplicity, we represent the whole deep cell population by a single cell type \mathcal{T}_{red} , and define each cell to exert *bipolar* protrusion forces over their \mathcal{T}_{red} neighbors. Parallel to the two polarization fields, we define two similar protrusion specifications (Fig. 8.57). When the polarization field originates from the diffusion of ligand Q_1 (resp. Q_2) secreted by the EVL (resp. YSL), protrusions are oriented along the axis \vec{U}_1 (resp. \vec{U}_2). In both cases, the intensity of the protrusive force f^A will be one of the explored parameters. We decided not to explore the influence of the passive force coefficients $w_{\mathcal{T}_{\text{red}} \mathcal{T}_{\text{red}}}^{\text{adh}}$ and w_{rep} , however, as they counterbalance the effect of f^A . In this study, both passive coefficients will be set to 1.



Figure 8.57: Active protrusion table A for the “Intercalation” case study. A: All the cells belonging to the \mathcal{T}_{red} population exert bipolar protrusions over each other along the polarization axis \vec{U}_1 derived from the gradient field of ligand Q_1 . B: The same cells can exert another bipolar protrusion along \vec{U}_2 derived from Q_2 . In both cases, the intensity of the protrusive force is f^A .

EVL model (imported from previous case study)

Finally, the last factor that we hypothesize to have an influence on the epibolic deformation of the embryo is the tangential tension in the EVL. As examined in the fourth case study (Section 8.4), the overall surface of the EVL clearly increases during epiboly. To reflect this, we have proposed a model of EVL that did not trigger the spreading intrinsically but rather resisted to it. We summarize here the contribution from this preceding case study to the present one.

Our model of cell proliferation and growth is controlled by various parameters:

EVL cells surface expand or shrink according to the external pressure exerted by their EVL neighbors. If the cell is compressed, i.e. the external pressure is greater than a positive threshold θ_E^+ , its lateral radius $R_i^{lat,E}$ decreases by a ratio γ_E . On the contrary, if the cell is under tension, i.e. the external pressure is smaller than a negative threshold θ_E^- , $R_i^{lat,E}$ increases by the same ratio γ_E . To control the resistance of the EVL against spreading, we modulate the EVL expansion threshold θ_E^- : if it is close to zero, any external tension will trigger expansion and potentially proliferation of the tissue, and the EVL will not resist spreading. If the absolute value of θ_E^- is high enough, then the EVL will not expand and proliferate, and it will resist spreading. Between these two extremes, we expect that the EVL will exhibit an intermediate degree of resistance, allowing us to decipher its influence on the whole epibolic motion.

Parameter space for the intercalation pattern in epiboly phase 1

In summary, the parameter space to be explored is 5-dimensional, and comprises:

- the type of polarization field: EVL-origin or YSL-origin
- the margin resistance $\theta_{m,\parallel}$
- the stochasticity of the polarization field controlled by λ_r
- the EVL resistance to external tension, controlled by θ_E^-
- the intensity of the protrusion force f^A

8.5.2 Simulation, Parameter Space and Validation

We envision two types of measures to characterize the epiboly phase 1 motion: *macroscopic measures* describing the whole embryo deformation, and *microscopic measures* describing the micro-level patterns of intercalation.

Macroscopic measures: temporal evolution of singular epiboly landmarks

Live specimen

While the *in toto* reconstruction of the zebrafish early development phenomenology (as described in Chapter 7) is a valuable strategy to decipher the multiscale interactions to contains, a qualitative understanding of the macro-scale deformation of the embryo is best achieved from Karlstrom and Kane’s 1996 “flipbook” of embryogenesis [315]. We derived from this movie a measurement of the macroscopic deformation occurring during epiboly phase 1 (see Movie S.5). We extracted 12 images from the oblong stage to the 50-percent epiboly stage (Fig. 8.58). We interpret from the analysis of these images that the embryo’s animal-vegetal (AV) axis is slightly away from the vertical, and that the timing between successive frames

(counted at the top left of each frame, first line) is very irregular. We decided to adjust this timing in hpf units (hours post fertilization) using the table provided in [315] (top left of each frame, second line). We manually annotated the most important landmarks on each image using 6 dots (see explanation in the figure caption).

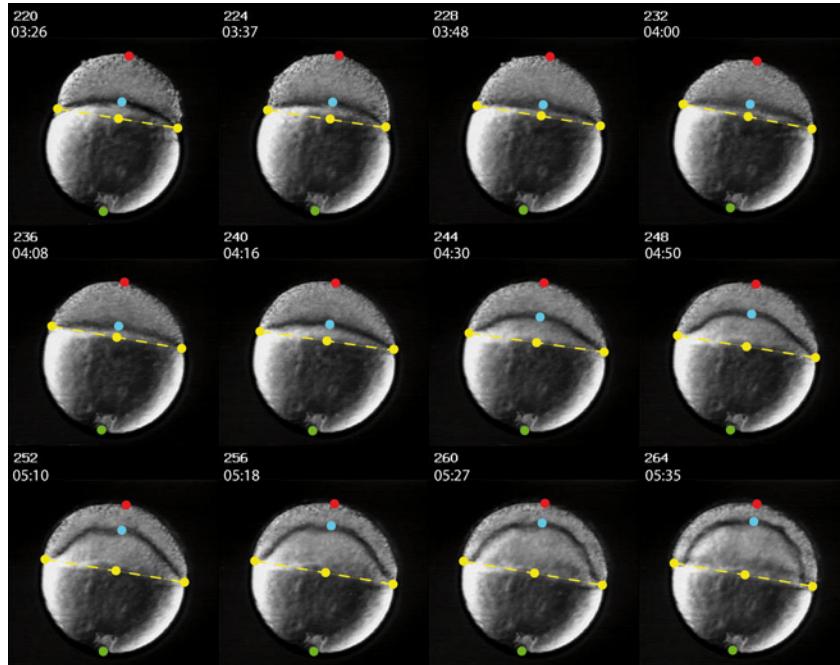


Figure 8.58: Macroscopic landmarks of the epibolic deformation from the “flipbook” specimen. Snapshots of the zebrafish development from the oblong stage to 50 percent epiboly extracted from the movie S.5 by Karlstrom and Kane [315]. Colored dots have been manually added to estimate the macroscopic morphological characteristics of the embryo: red dots signal the animal pole of each embryo, green dots the vegetal pole, blue dots the animal-pole limit of the yolk, and triplets of yellow dots delineate the margin, where the left and right dots identify the external position of the margin and the central one correspond to their averaged projection on the AV axis. The time value displayed below the image id is the time in hours post fertilization given by [315]. These timings do not scale linearly with the image ids and have been renormalized.

The macroscopic measures were inferred from the singular landmarks described above and consisted in the temporal evolution of four absolute distances (Fig. 8.59, dashed lines):

- the embryo height from the vegetal pole to animal pole (red dashed line)
- the margin height from the vegetal pole to the central marginal position (green

dashed line)

- the yolk height from the vegetal pole to the yolk animal pole (blue dashed line)
- the margin width from the left to the right marginal positions (yellow dashed line)

These measures provide an absolute macroscopic description of the deformation occurring during phase 1 epiboly. However, in order to compare different embryos, either live or simulated, a normalization of the distances is required. This normalization implies a loss of part of the information. We chose to use the embryo height as a reference for the comparison, leaving three remaining measures:

- the normalized margin height H_m^e , obtained by dividing the margin height distance by the embryo height (green solid line): it characterizes the overall covering of the yolk by the cells; this measure will tend to zero as the tail bud closure proceeds./li>
- the normalized yolk height H_Y^e , obtained by dividing the yolk height by the embryo height (blue solid line): it characterizes the doming of the blastoderm and the bulging of the yolk
- the embryo width to height ratio or *sphericity* coefficient C_Φ^e , obtained by dividing the margin width by the embryo height (yellow solid line): it characterizes the sphericity of the whole embryo; this value would be equal to one for a spherical embryo, less than one if the embryo is elongated along the AV axis, and more than one if the embryo flattens along the AV axis.

These simple measures allow to distinguish two macroscopic phases of deformation (grey areas in Fig. 8.59): first, the sphere transition deformation, occurring between 3h37min (hpf) and 4h00min (hpf) in the flipbook, and second, the doming deformation, starting 8 min later and going on between 4h08min (hpf) and 5h35min (hpf). During the sphere transition, the absolute height of the embryo decreases, while its sphericity increases rapidly. The flattening at the sphere stage of the yolk cell-blastoderm interface appears moderate, and it is possible that other specimens behave somewhat differently in this respect (as in Fig. 8.54C for example). Additional embryos would be useful to refine this measure (and the other measures as well). The doming transition is accompanied by an important move of the blastoderm margin toward the vegetal pole and an even more important relative displacement of the yolk cell's animal pole, while the overall sphericity slowly increases.

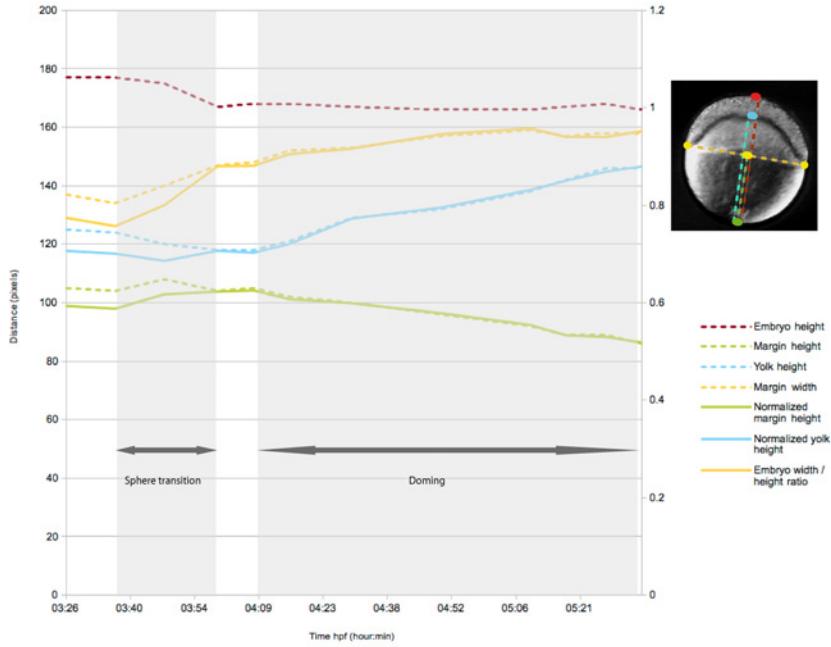


Figure 8.59: Macroscopic measures of the epibolic deformation in the “flipbook” specimen. The measures defined by the macroscopic landmarks displayed in Fig. 8.58 are shown. The red line is the plot of the distance between the animal pole (AP) of the embryo and the vegetal pole (VP) of the yolk (embryo height). The green line is the plot of the distance between the projection of the margin on the animal-vegetal (AV) axis and the VP of the yolk (margin height). The blue line is distance between the AP of the yolk and the VP of the yolk (yolk height). The yellow line is the lateral distance between the margin positions (margin diameter). The dashed lines give the absolute distance between landmarks in pixels (left ordinate axis). The continuous lines give the normalized distances (right ordinate axis). The normalization is obtained by dividing each value by the current yolk height (i.e. dashed red line). The abscissa gives the time in hour post fertilization.

Simulated specimen

In contrast to the live measures that were performed on 2D images, the simulated embryo is a 3D structure that requires an adaptation of the automated measurements. All the measures were performed in reference to the AV axis and a point \mathbf{M}_o in the center of the embryo (we use here upright boldface notation for all 3D point coordinates and vectors). In the simulation, the AV axis unitary vector \mathbf{U}_{AV} is specified a priori in the initial conditions (Fig. 8.60). The embryo's center \mathbf{M}_o is calculated and updated at each simulation time step by averaging the positions of the yolk membrane (ym) and the yolk interior (yi) particle:

$$\mathbf{M}_o = \frac{1}{N_{yi} + N_{ym}} \left(\sum_{i=0}^{N_{yi}-1} \mathbf{X}_{yi,i} + \sum_{i=0}^{N_{ym}-1} \mathbf{X}_{ym,i} \right)$$

where $\mathbf{X}_{yi,i}$ (resp. $\mathbf{X}_{ym,i}$) is the position of the yolk interior (resp. yolk membrane) particle i and N_{yi} (resp. N_{ym}) is the total number of yolk interior (resp. yolk membrane) particles.

The embryo vegetal pole \mathbf{M}_V^E , embryo animal pole \mathbf{M}_A^E and yolk animal pole \mathbf{M}_A^Y are obtained by calculating the scalar product between the AV axis \mathbf{U}_{AV} and each ym or EVL particle's relative position with respect to the center, then selecting the position that realizes the maximum (farthest value in the positive direction) or minimum (farthest value in the negative direction) of this product:

$$\begin{aligned} \mathbf{M}_V^E &= \arg \min_{\mathbf{X}_{ym,i} \in \mathcal{S}_{ym}} ((\mathbf{X}_{ym,i} - \mathbf{M}_o) \cdot \mathbf{U}_{AV}) \\ \mathbf{M}_A^E &= \arg \max_{\mathbf{X}_{E,i} \in \mathcal{S}_E} ((\mathbf{X}_{E,i} - \mathbf{M}_o) \cdot \mathbf{U}_{AV}) \\ \mathbf{M}_A^Y &= \arg \max_{\mathbf{X}_{ym,i} \in \mathcal{S}_{ym}} ((\mathbf{X}_{ym,i} - \mathbf{M}_o) \cdot \mathbf{U}_{AV}) \end{aligned}$$

where \mathcal{S}_{ym} is the set of yolk membrane particles' position and \mathcal{S}_E is the set of EVL cell particles' position. The projection of the margin on the AV axis \mathbf{M}_m is obtained by averaging the scalar projection of all margin yolk membrane (MYM) particles' relative positions on the AV axis:

$$\mathbf{M}_m = \mathbf{M}_o + \frac{1}{N_m} \sum_{i=0}^{N_m-1} ((\mathbf{X}_{m,i} - \mathbf{M}_o) \cdot \mathbf{U}_{AV}) \mathbf{U}_{AV}$$

where $\mathbf{X}_{m,i}$ is the position of the MYM particle i and N_m is the number of MYM particles. Finally, the embryo's width W is twice the average of the radii defined by the distances between the MYM particle positions and their common projection \mathbf{M}_m :

$$W = \frac{2}{N_m} \sum_{i=0}^{N_m-1} \|\mathbf{X}_{m,i} - \mathbf{M}_m\|$$

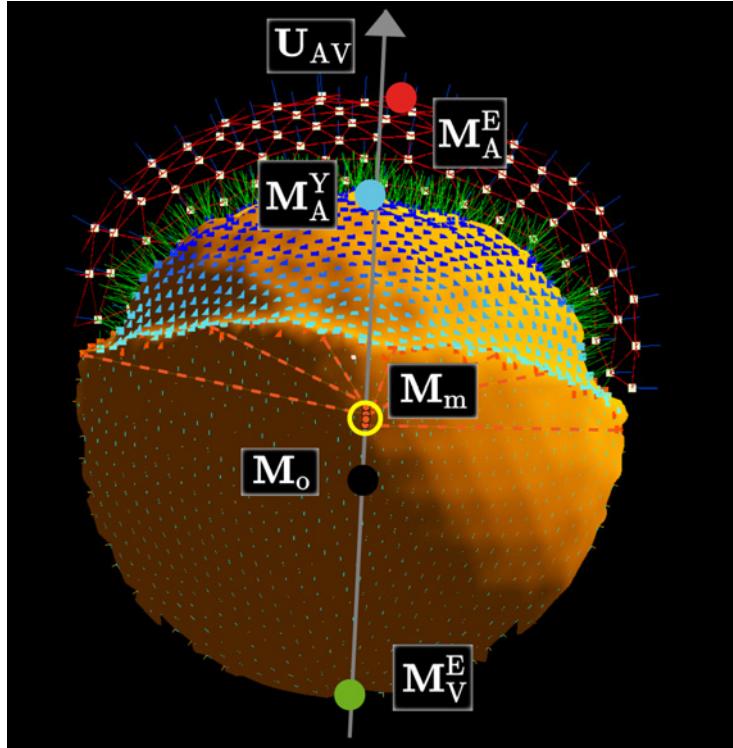


Figure 8.60: Macroscopic landmarks of the epibolic deformation in simulated specimens. Landmark dots have been manually added for the purpose of the illustration. They represent the macroscopic landmarks automatically calculated by the simulated measurements. The grey line is the animal-vegetal (AV) axis, *a priori* specified for the initialization. The red dot is the embryo animal pole M_A^E , the blue dot is the yolk animal pole M_A^Y , the black dot is the embryo center M_o , the green dot is the embryo vegetal pole M_V^E , the orange dot is the margin “level” M_m , i.e. the average of all the projections of margin yolk membrane particles (small orange circles) on the AV axis (only a few of these particles are displayed). Note that the position of the yolk animal pole, embryo animal pole and embryo vegetal pole are never exactly aligned on the AV axis as they must coincide with one particle’s position. The deep cells are not displayed here, leaving a carved out domain between the yolk cell and the EVL.

Similarly to the live specimen, we define the normalized margin height H_m^s , the normalized yolk height H_Y^s and the sphericity ratio C_Φ^s as the following:

$$H_m^s = \frac{\|M_m - M_V^E\|}{\|M_A^E - M_V^E\|}, \quad H_Y^s = \frac{\|M_A^Y - M_V^E\|}{\|M_A^E - M_V^E\|}, \quad C_\Phi^s = \frac{W}{\|M_A^E - M_V^E\|}$$

The results of these macroscopic measures of epibolic deformation on the simulated embryo are shown in Fig. 8.61 and Movie 8.11.

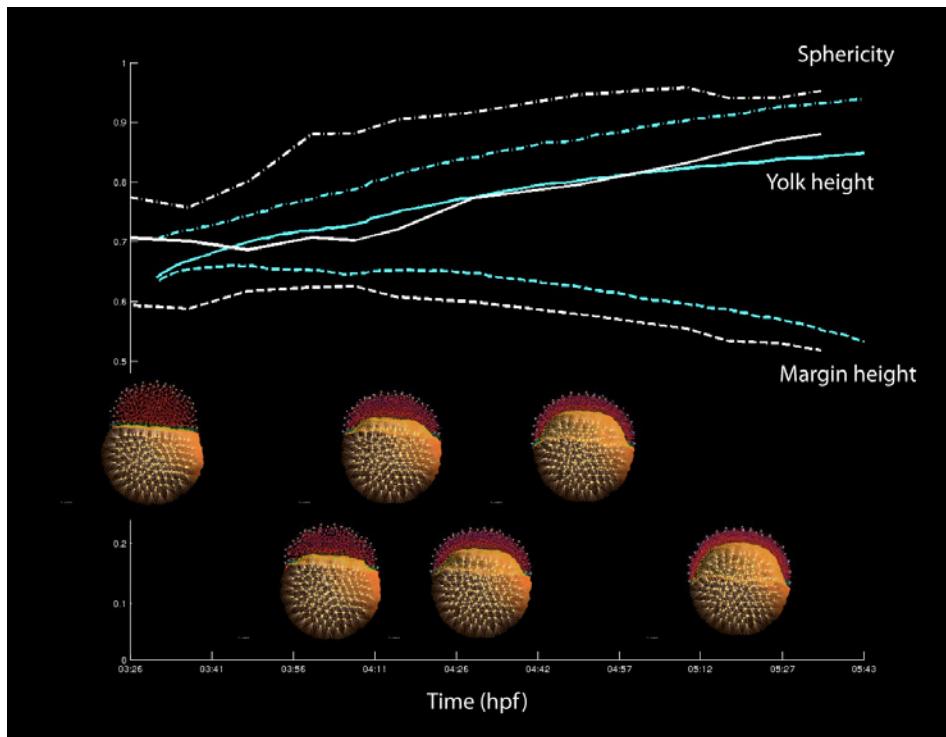


Figure 8.61: Macroscopic measures of the epibolic deformation in a simulated specimen. Six snapshots corresponding to different stages of a simulated embryo are displayed along with the temporal evolution of the corresponding macroscopic measures (blue), compared with the flipbook measures (white). Sphericity is represented by the dash-dot lines, normalized yolk height by the solid lines and normalized margin height by the dashed lines. The embryos' positions correspond approximately to the snapshot times in abscissa. The simulated embryo parameters are: ligand source Q_2 on the yolk, $\theta_{m,\parallel} = 5.33$, $\lambda_r = 0$, $\theta_E^- = 56.67$ and $f^A = 3556$. The movie from which these snapshots are extracted can be viewed in 8.11.

Our first observations from these preliminary results is that the deep cells' active protrusive behavior seems sufficient to drive the embryo deformation during epiboly phase 1, in the context of the simulation. As it can be seen on Fig. 8.61, the temporal evolution of the macroscopic measures fits reasonably well the data extracted from Karlstrom and Kane's flipbook. The main discrepancy with the flipbook is the apparent lack of sphere transition leading to the sphere morphology during the first hour. We do not observe the characteristic phase of increased sphericity concomitant with a *statu quo* for the normalized yolk and margin heights. In the simulated

embryo's trajectory, the early increase of sphericity is always simultaneous with the decrease of at least one of the other macroscopic measures. A possible explanation could be that the initialization of the simulation is not close enough to the real high stage, with a yolk-deep cell interface already presenting a flat shape in the simulation.

After this study of one particular embryo under a fixed set of parameters, we now proceed to a more in-depth exploration of the space parameter.

Fitness function

Aided by the above measures that perform an automated quantitative comparison between live and simulated embryos, we designed three objective subfunctions to evaluate the similarity of the simulated embryos with the flipbook specimen:

- F_m , the normalized margin height objective function,
- F_Y , the normalized yolk height objective function,
- F_Φ , the sphericity objective function.

The three objectives functions have the same expression:

$$F_m = \sum_{t=t_1}^{t_{12}} |H_m^s(t) - H_m^e(t)|, \quad F_Y = \sum_{t=t_1}^{t_{12}} |H_Y^s(t) - H_Y^e(t)|, \quad F_\Phi = \sum_{t=t_1}^{t_{12}} |C_\Phi^s(t) - C_\Phi^e(t)|$$

where $t_1 \dots t_{12}$ are the timings of the 12 images extracted from the flipbook specimen (Fig. 8.58). These three objectives functions are then merged into a global fitness function F via a simple averaging operation:

$$F = \frac{1}{3}(F_m + F_Y + F_\Phi)$$

In contrast to Section 8.2.2, the contribution of each objective function is added without normalization as these functions are all of the same kind, i.e. differences of normalized measures.

Fitness landscape

We explored the parameter space of the simulation to find the best match with the live measurements. This space is 5D, comprising the source of ligand diffusion (EVL or YSL), the margin resistance $\theta_{m,\parallel}$, the stochasticity of the polarization field controlled by λ_r , the EVL resistance to external tension controlled by θ_E^- , and the intensity of the protrusion force f^A . The parameter space was regularly sampled over the following range with the following cardinalities (Fig. 8.62):

A number of insights could be gained from the fitness landscape analysis. A general trend is that the protrusive force intensity and the randomness coefficient have

	Min.	Max.	Cardinality	Unit
source	1	2	2	-
$\theta_{\text{mym},\parallel}$	1	40	10	N
λ_r	0	0.1	10	-
θ_{evl}^-	10	150	10	N.m^{-2}
f^A	0	8000	10	N

Figure 8.62: Range, cardinalities and units of the five parameters explored in this study.

counterbalancing effects. A higher protrusive force coupled with a higher random factor produces a fitness value similar to a couple of factors with lower values, as indicated by the isolines of Figs. 8.63-8.65. The profile of the isolines gives a relationship between both parameters. It appears to be supralinear as counterbalancing an increase of the protrusion force intensity requires an exponential increase of the randomness parameter.

For couples of parameters (f^A, λ_r) situated below the isoline passing through the point of coordinates (2000, 0.05), we do not observe any macroscopic epiboly behavior. We may however qualitatively distinguish different microscopic behaviors in this area: for low levels of protrusive force intensity, the lack of epibolic deformation is due to the lack of intercalating behavior at the cellular level. On the contrary, for high levels of protrusive force intensity and randomness, the cells start intercalating inefficiently, with cells sliding on each other in a fluid-like manner, and tissue cohesion is lost (Movie S.6). This behavior is due to an inequality between the active and passive forces in favor of the protrusive forces. Further study of these mechanisms would require a preliminary calibration between both types of forces to ensure that these behaviors do not occur. In the following, we performed an *a posteriori* calibration of parameter f^A at value 3555.

Eliminating the f^A dimension allows us to use the 4D visualization tool introduced in Section 8.1.2: the results can be seen in Figs. 8.66-8.69. This tool and the complete fitness datasets are available here for download (MATLAB required). The comparison of the fitness landscapes associated with the two polarization scenarii (sources 1 and 2) points out their similarities and differences, as discussed below.

A general trend is the penalizing effect of the polarization field randomness: all objective functions perform poorly in this domain (top parts of all 3D charts). The only exception concerns the normalized margin height measure F_m , which tempers this effect when conjugated with low EVL resistance to tension θ_E^- , and low margin resistance to deep cells' pressure $\theta_{\text{m},\parallel}$ (Fig. 8.66). Observations of the simulated phenotypes in this particular domain (Movie S.8) show that the relaxed state of the EVL and margin, coupled with the low efficiency of the deep cells' protrusive activity (due to high randomness), produce a slight move of the margin toward the vegetal pole along with a de-flattening of the yolk-deep cells interface.

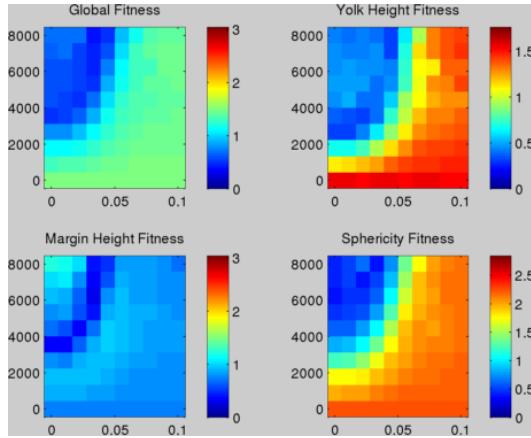


Figure 8.63: Fitness landscapes as a function of the random parameter λ_r and the protrusive force intensity parameter f^A . Top Left: global fitness function. Top Right: normalized yolk height objective function F_Y . Bottom Left: normalized margin height function F_m . Bottom Right: sphericity objective function F_Φ . On each plot, the random parameter λ_r is shown in abscissa and the protrusive force intensity f^A in ordinate. The color maps scale between zero and the maximum value of the fitness landscape that they are associated with (to the left).

This observation emphasizes the importance of the embryo's external tension (EVL+Margin) for the shaping of the cellular domain (see third case study, "Shaping the Blastula", in Section 8.3). We also observe that in the scenario of the EVL as a ligand source (value 1), the EVL's topology becomes more irregular. A feedback coupling involves the perturbation of the polarization field from the disorganized EVL, disorienting the deep cells, which in turn reinforces the EVL disorganization. Both the normalized yolk height and the sphericity measures perform badly as yolk bulging does not occur.

At the opposite side of the parameter spectrum, i.e. with low polarization randomness λ_r and high EVL and margin resistance, the deep cells' intercalation gains in efficiency but is constrained by the margin lock, preventing them to spread over the yolk. Interestingly, the two polarization scenarios offer alternative behaviors in response to this abnormal condition. In the YSL-based field, the deep cells' intercalation exerts a spreading force that deforms the overlying EVL. This deformation is actually much more important in the EVL-based field, where the epiboly's spreading triggers the blastoderm's wrenching from the yolk surface (Movie S.9). This pathological behavior is not expected and can be explained by an inadequacy of the intensity of the forces, without the random axis-induced damping. Obviously, both the sphericity and the normalized yolk height measures penalize the EVL-based scenario, compared to the YSL one (Fig. 8.68 and Fig. 8.67, bottom right corner of

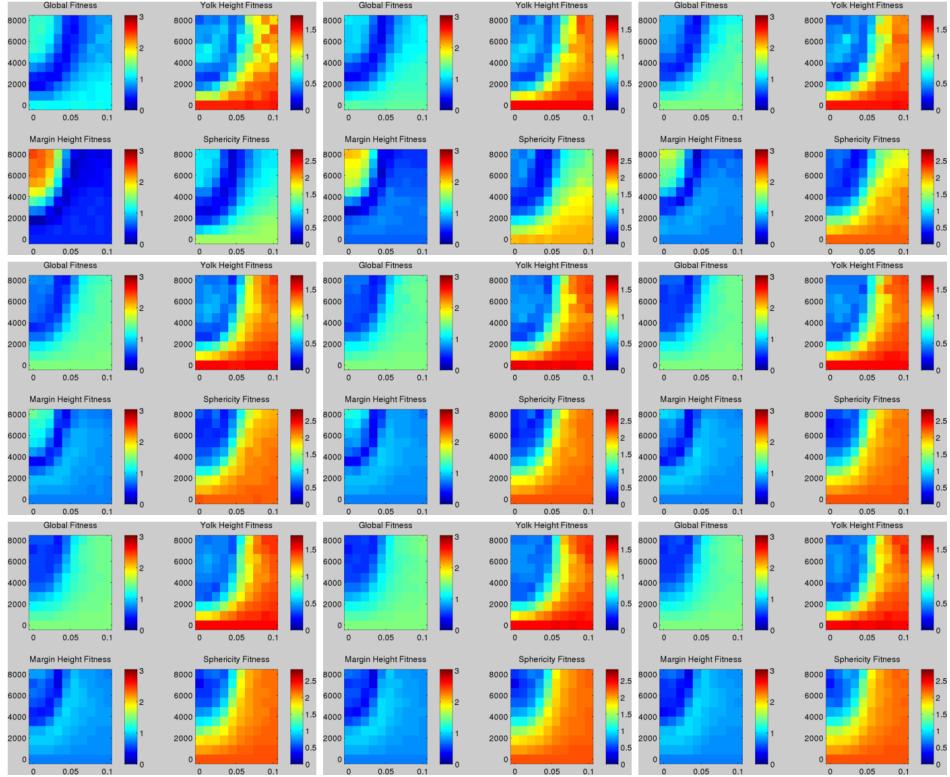


Figure 8.64: Fitness landscapes as a function of the random parameter λ_r and the protrusive force intensity parameter f^A , with the yolk as a ligand source (value 2). Each of the 9 subplots is equivalent to the landscapes presented in Fig. 8.63. From top left to bottom right, the margin resistance $\theta_{m,\parallel}$ increases regularly from 1 to 35.666 (the last subplot for value 40 was removed to avoid excessive image reduction). The threshold controlling the EVL resistance θ_E^- is set to 72. The whole dataset of these fitness landscapes is available here.

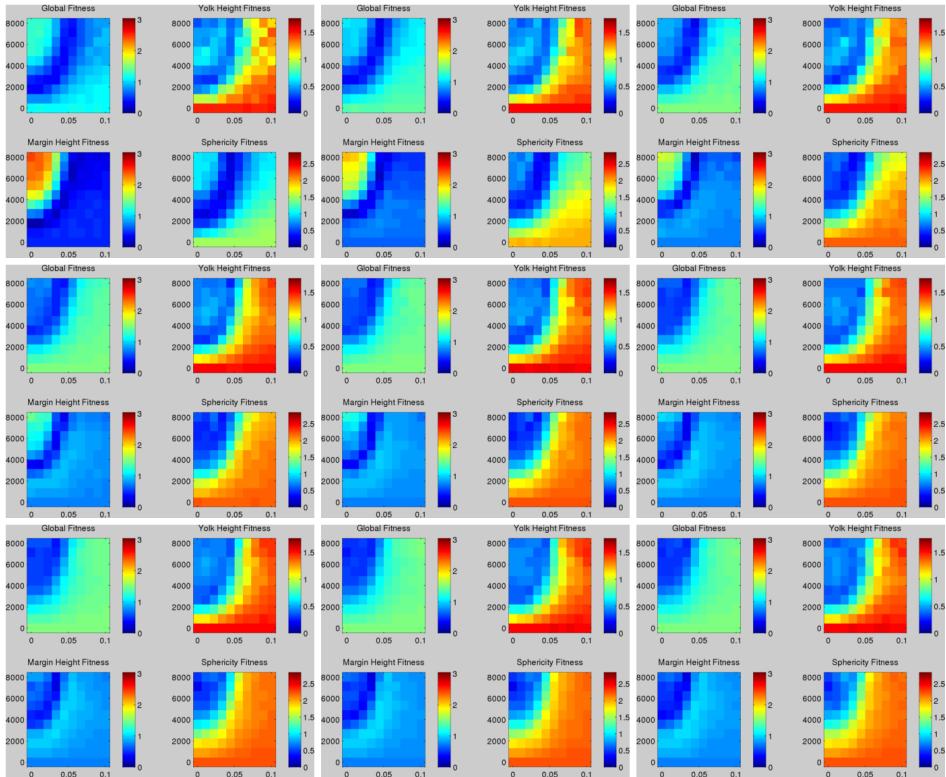


Figure 8.65: Fitness landscapes as a function of the random parameter λ_r and the protrusive force intensity parameter f^A , with the EVL as a ligand source (value 1). Except for the source, the parameter space mapped here is the same as in Fig. 8.64

the cubes).

Another dramatic scenario is obtained if an efficient intercalation (no randomness in the field) is happening together with strong EVL resistance and a weak margin resistance (Movie S.7). The previous scenario happens again, i.e. the EVL blocks the spreading of the deep cells, except that at some point the margin still receives pressure that makes it move toward the vegetal pole, pulling the yolk membrane toward the animal pole. The accumulating yolk membrane ends up being rolled in a unrealistic fashion, stretching the inner yolk membrane particles, and finally allowing the strongly intercalating deep cells to perforate the yolk membrane, and penetrate into the yolk. As the spatial neighborhood algorithm is not designed to deal with this situation, a flow of deep cells enters the yolk.

The last example of aberrant development was detected by the normalized margin measure, due to its too-vegetal margin position (Fig. 8.66, bottom right corner), but both the normalized height and the sphericity measures evaluated the developmental trajectory of the embryo as excellent (Fig. 8.66, bottom right corner). This mischaracterization of an abnormal behavior highlights the difficulty to *a priori* envision the potential behaviors of a simulated model.

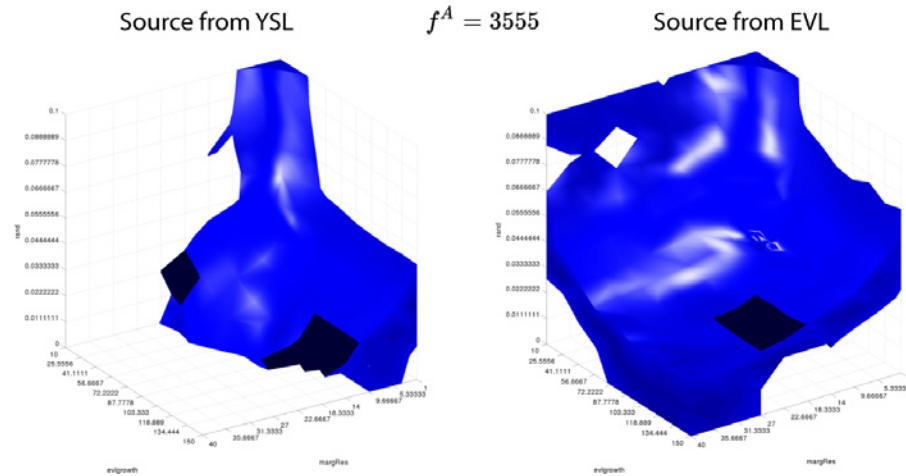


Figure 8.66: 3D plot of the normalized margin height objective function F_m . The isosurface represents the best fitness volume (threshold 0.38). The protrusive force intensity f^A has been set to 3555. The vertical axis indicates the random parameter λ_r , the depth axis indicates the threshold controlling the EVL resistance θ_E^- and the horizontal axis indicates the margin resistance threshold $\theta_{m,\parallel}$. This 3D objective function landscape has been rotated, unlike the following ones, to show a better angle of view. The interactive 4D visualization tool is available here for download (MATLAB required)

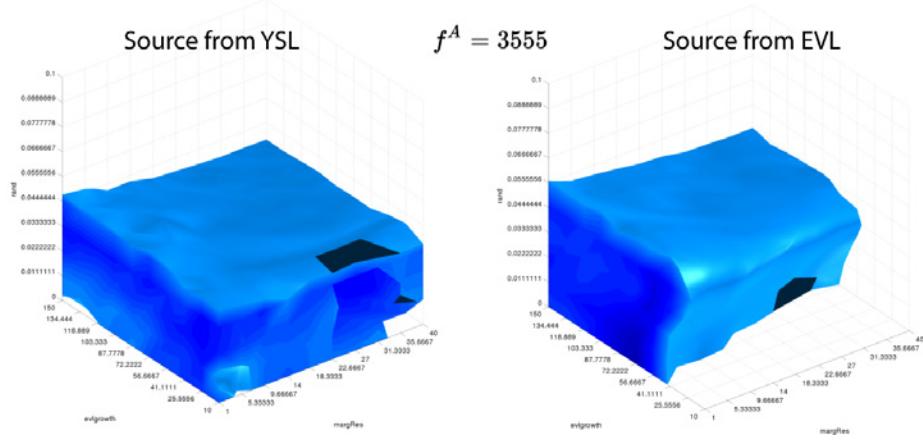


Figure 8.67: 3D plot of the normalized yolk height objective function F_Y . The threshold of the isosurface is set to 0.56. See Fig. 8.66 for details.

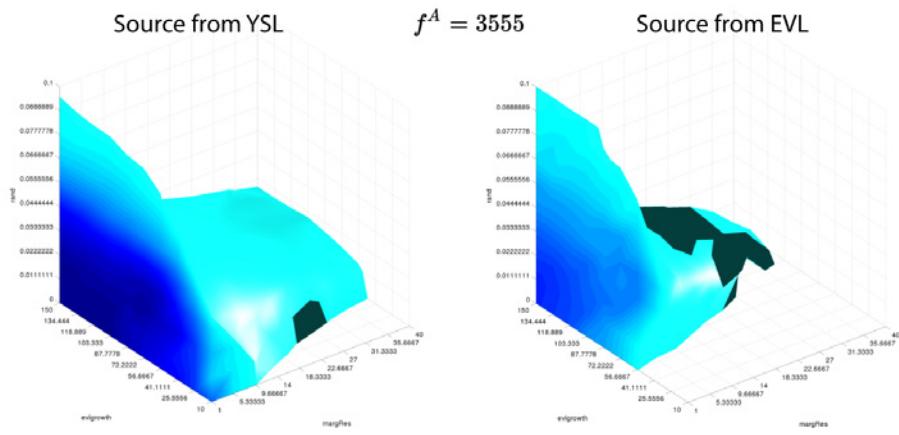


Figure 8.68: 3D plot of the sphericity objective function F_Φ . The threshold of the isosurface is set to 1.20. See Fig. 8.66 for details.

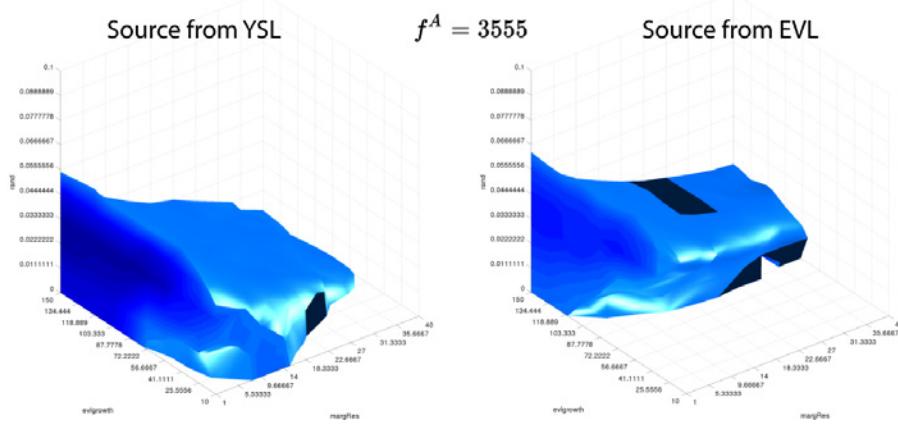


Figure 8.69: 3D plot of the global fitness function F . The threshold of the isosurface is set to 0.69. See Fig. 8.66 for details.

Microscopic measures: temporal evolution of cells local neighborhood

While the above measures give a valuable estimation of the *in toto* macroscopic behavior, they do not permit to evaluate the micro-scale of cell interactions during the process. The rapid flattening of the deep cells' mass, from 12 cell rows at the sphere stage to about 4 cell rows at the germ ring stage, must involve an intercalation of the cells along the embryo's radial axes. We assessed cell behaviors during epiboly phase 1 by focusing on their radial displacement. By "radial displacement", we mean the temporal evolution of the relative position of the cells in the depth of the cell mass. We denote here by \mathcal{L}_i^R the radial line passing through the center of the embryo \mathbf{M}_o and a deep cell i at position X_i . On this line, we denote by $\rho_i^R(t)$ the "normalized" position of cell i between the EVL- \mathcal{L}_i^R intersection point $X_{i,E}$ and the YSL- \mathcal{L}_i^R intersection point $X_{i,Y}$ at time t (Fig. 8.70):

$$\rho_i^R(t) = \frac{X_i(t) - X_{i,Y}(t)}{X_{i,E}(t) - X_{i,Y}(t)}$$

With these notations, we now define the microscopic measure $M_i^R(t_0, t_n)$ characterizing the normalized radial displacement of a cell i during a time interval $[t_0..t_n]$ as the sum of the absolute values of its relative displacements at each time step:

$$M_i^R(t_0, t_n) = \sum_{k=0}^{n-1} |\rho_i^R(t_{k+1}) - \rho_i^R(t_k)|$$

In a minimal rearrangement scenario with a purely increasing (or decreasing) ρ_i^R value, this measure boils down to the difference between the first and last relative radial positions: $M_i^R(t_0, t_n) = \rho_i^R(t_n) - \rho_i^R(t_0)$. However, if cells are moving back and forth radially, then this measure takes on higher values (Fig. 8.71).

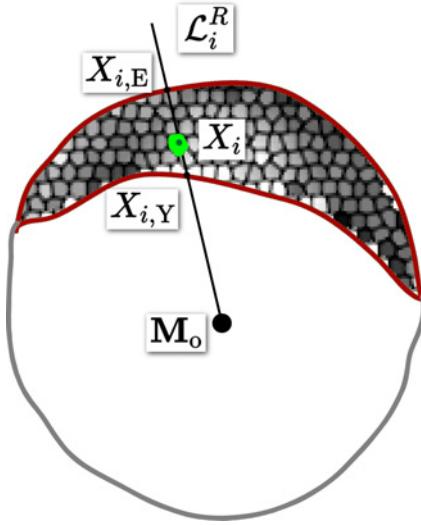


Figure 8.70: Normalized position of the cells along the radial axis \mathcal{L}_i^R . The measure is applied to the cell i in green shape with its approximate center as a black dot, at position X_i . It represents the fraction of the height of X_i on the segment going from $X_{i,Y}$ to $X_{i,E}$.

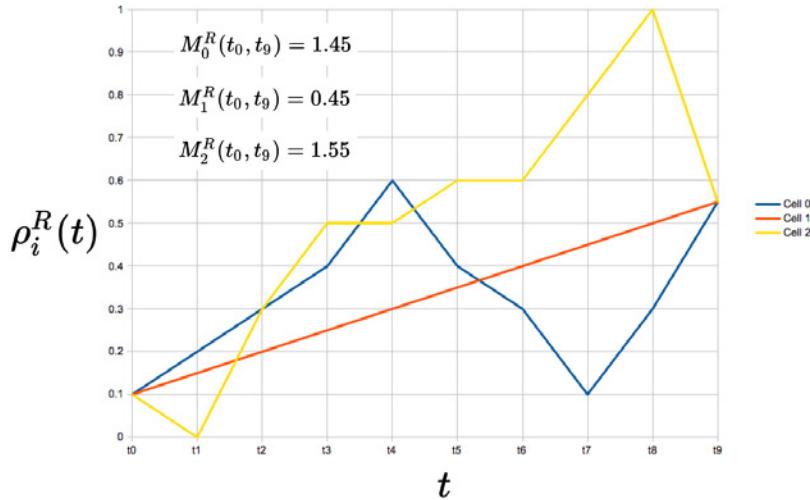


Figure 8.71: Illustration of the microscopic measure of cell intercalation M_i^R . Three fictive cell trajectories in normalized radial displacement space: Cell 0 (blue), Cell 1 (red), Cell 2 (yellow). Cell 1 performs a continuously increasing radial displacement between position 0.1 at t_0 and position 0.55 at t_9 , yielding a measure of $M_1^R = 0.45$. Cells 0 and 2 display random radial displacements through time, which are characterized by higher microscopic measure values $M_0^R = 1.45$ and $M_2^R = 1.55$.

Another kind of microscopic characterization of the intercalation patterns deals with the dispersion along the tangential plane of the embryo. We designed a measure of this dispersion relying on the temporal evolution of a population originating from a radial cylinder (Movie 8.12). This measure is not explained here. To be used, it would need to be calculated at different positions along the animal-vegetal axis. A possible refinement of this cylinder measure would consist of integrating a vertical stack of colored layers in order to measure more precisely the patterns of radial intercalation and their statistical repeatability.

8.5.3 Discussion

Our simulations validate the fact that individual cell protrusive activity is sufficient to drive phase 1 epiboly. We have pointed out, however, the absence of sphere stage transition. In the simulated embryo's trajectories, early increasing sphericity always goes with a progression of the embryo margin toward the yolk. We proposed as an explanation that the initialization of the simulation is not close enough to the real high stage, with a yolk-deep cell interface already presenting a flat shape in the simulation.

Another explanation is that an additional mechanism is missing in the explored parameter space, either via a different organization of the polarization field, or via new behavioral rules in the different compartments of the embryo. A refinement of the mechanical properties of the YSL (Yolk Syncytial Layer), which are homogeneous in the current state of the model, would be an interesting improvement, as the YSL must have a different behavior than the YCL (Yolk Cytoplasmic Layer). Such a refinement might also help achieve the embryo doming. We can imagine that an internal YSL structure bears an external tension at the margin until a certain limit, beyond which its structure breaks, leading the doming deformation.

Exploring the parameter space does not only aim at finding the “best” sets of parameters, or discuss the relationship between the parameters. It is also an exploration of the limits of the model and its implementation. Three levels of viability are accessible in this kind of study: the domain of viability expressed by the measures (described via the fitness landscapes), the domain of unmeasured viability (aberrations that perform well in the fitness landscapes but circumvent their criteria), and the domain of numerical simulation viability, which restricts the accessibility to the other ones. The latter may be extended by modifying the algorithmic procedure used to compute the model *in silico*, for example by reducing the simulation time step, or using higher order integration schemes. The unmeasured viability, for its part, constitutes a peculiar artifact of this type of simulations, which evolve in high dimensional spaces that are not fully mastered (due to both technical and cognitive limitations). An example of further investigation that would allow us to characterize the epiboly phase 1 episode and apprehend the effective properties of cell motility is to measure the epibolic behaviors at the cellular level. Such measures

are currently investigated, both *in silico* and *in vivo*. We provided at the end of Section 8.5.2 the main lines of what we envisioned at the microscopic level.

8.6 Gastrulation

Gastrulation starts with the internalization of the hypoblast. This is usually said to happen at the shield stage although there are signs at the germ-ring stage (5.7 hpf). Gastrulation separates the blastoderm in three concentric cell layers: the endoderm lying on the yolk, the mesoderm above, and the ectoderm forming the epiblast which does not internalize and maintain a contact with the EVL. Shortly after the beginning of the internalization (30 min [277]), the cells belonging to all three layers start converging toward the dorsal side of the embryo and thicken the embryonic axis in the antero-posterior and radial directions via local intercalation (Convergence-Extension, CE). The conjugated motion of all layer leads to the closing of the tail bud at 10h.

Multiple mechanisms have been hypothesized to generate the complex deformations occurring over the time span of these 4 hours (see our review of gastrulation in Section 6.3). In this case study, we hypothesize that the protrusive activity discussed in previous sections could be sufficient to lead to gastrulation morphogenetic steps, with a polarization field oriented toward the dorsal axis. Specifically, we aim at examining three particular mechanisms involved in the gastrula morphogenesis:

- a. the contribution of the orientation of cell division to CE
- b. the mechanical interplay between the hypoblast deformation and the epiblast deformation
- c. the role of the polarization field in driving CE
- d. the role of internalization in the late epibolic motion.

a. Influence of the division orientation axis on convergence and extension

The convergence and extension tissue deformation is a generic morphogenetic process, which plays a major role in vertebrate embryogenesis. Various cellular mechanisms have been proposed for the elongation of tissue, and they can be classified in two categories: an active *cell deformation* in the direction of elongation with low rearrangement of the local cell neighborhood (*C. elegans*, ascidians such as *Ciona* or *Phallusia*); and an active *cell rearrangement* without cell elongation in the direction of the tissue elongation (ascidians such as *Ciona* or *Pallusia*, *Drosophila*). Another independent mechanism that has been demonstrated to participate in tissue elongation is the *orientation of division* along the elongation axis (avians [316], *C. elegans*, *Drosophila*). Gong et al. [291] identified the controlled cell division orientation as a driving force for axis elongation during the zebrafish gastrulation.

In their study, they use *in vivo* confocal imaging showing that mitotic divisions in the dorsal region of the gastrula are preferentially oriented along the animal-vegetal axis. They show that, in this region (dorsal equatorial), the divisions are not oriented along the (medio-lateral) axis of cell elongation (contrary to the postulate of Hertwig in 1893) but rather along the internal polarization axis of the non-canonical Wnt pathway (also denominated planar cell polarity pathway, PCP). By perturbing the function of a PCP-related molecule, Dishevelled (Dsh), they observe the conjugated randomization of the axis of division and the disruption of the convergence-extension deformation. They conclude that division along the dorso-ventral axis favors convergence-extension.

However, it is not clear that the link between division alignment and CE is correlative or causal. We believe that the impossibility to separate the randomization of the division orientation and the other PCP related processes prevents the definitive assessment of this causal link. We hypothesize that the cells' protrusive behavior is responsible for the CE deformation and that, in addition to division axes, the PCP perturbation is also randomizing the orientation of protrusions. This hypothesis would relegate cell division orientation to creating CE as a side effect rather than being its driving force. To test this hypothesis, we introduce two independent randomization factors, one for the associated field driving polarized protrusions, and the other acting on cell division orientation. The possibility to decouple the two mechanisms allows us to test different scenarios:

- i. whether CE is disrupted when division axis is randomized while the protrusive activities are not
- ii. whether, when the protrusive activities are directed by a PCP-related mechanism, division orientations act as a reinforcing factor of the elongation
- iii. whether a randomized division axis can be reoriented along the antero-posterior axis by the action of the neighboring cells' protrusive behavior.

This study would require the design of a macroscopic measure of the CE deformation. This measure could be the evolution of the length of the antero-posterior and medio-lateral axes through time in the different above scenarios.

b. The mechanical interplay between the hypoblast deformation and the epiblast deformation

The study of (*M*)*Zoep* mutants demonstrate a certain form of autonomy in CE deformation of the different layers of the zebrafish gastrula. An *oep* mutant is missing a cofactor of the receptor of the Nodal pathway, *One-eye-pinhead* (*OEP*), leading to the absence of activation of the downstream targets of the pathway (e.g. *goosecoid*, *sox17*...). The *zoep* mutant lacks endoderm and prechordal plate. *MZoep* is

devoid of both the maternal and the zygotic component of *oep*, and lacks endomesoderm except for a small population of ventral mesodermal cells giving rise to a few somites in the tail. Despite these severe perturbations of the layers, factors establishing dorso-ventral asymmetry still exist in the ectoderm (Fig. 8.72B), and this layer undergoes CE. The antero-posterior patterning of the neural plate roughly maintains its organization (Fig. 8.72A). The potential relative mechanical autonomy of the epiblast raises the question of the influence of the hypoblast on the epiblast convergence. It has been suggested that the hypoblast migrates on the internal surface of the epiblast. We hypothesize that such an inter-layer interaction might slow down the convergence speed of the epiblast and accelerate the speed of hypoblast convergence. We aim at testing this hypothesis *in silico* to decipher the mechanical interaction between the two layers. This hypothesis would be confirmed if the epiblast would converge faster in the absence of hypoblast. This is actually a possible interpretation of the *MZoep* mutant phenotype although the interdependence of nodal activation, downstream targets expression patterns, convergence and extension makes it difficult to draw definitive conclusions (Fig. 8.72B).

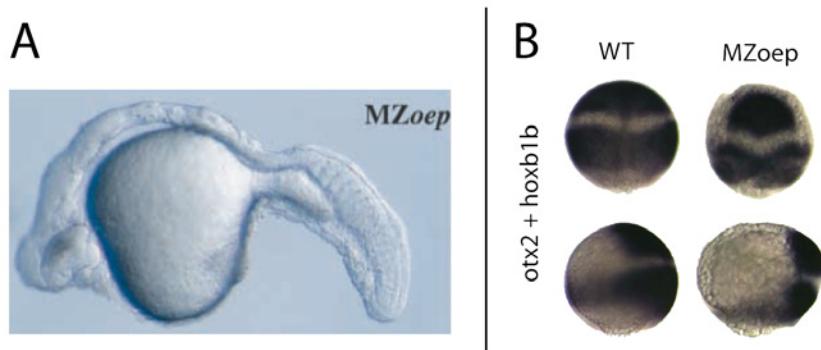


Figure 8.72: *MZoep* mutant. A: Live *MZoep* mutant embryo at 30 hours of development. The ventral mesoderm gave rise to a few somites in the most posterior part of the embryonic axis. The epiblast converged, shaping the antero-posterior axis. Image and caption adapted from Mathieu et al. [317] B: Dorsal markers expressed in the ectoderm (*otx2*, anteriorly, and *hoxb1b*, posteriorly), are still present in the *MZoep* mutant. Both domains are however narrower in the mutant. An interesting interpretation would be that the narrower domains are the result of a faster convergence in the mutant, due to the lack of hypoblast. Image adapted from Jia et al. [318].

c. Assessing the polarization field hypothesis in driving CE

We could test *in silico* the “source” of the polarization field that we apply to drive CE. In the *Zoep* mutant, the antero-posterior elongation is much less important than

in wild-type. The absence of prechordal plate correlates with the absence of ventral brain and might perturb the expression of the polarization inducing factors in both the epiblast and the hypoblast. We hypothesized that the *oep* mutant phenotypes can be reproduced by adjusting the specification of the polarization field in the simulations.

d. The role of internalization in the late epibolic motion

As mentioned in the review of the possible mechanisms underlying the epibolic motion occurring after 70-percent epiboly, we proposed that cells may be internalizing until the closure of the tail bud (Section 6.3.1). This possibility would lead us to reconsider the scenario proposed by Keller et al. [277], i.e. the absence of internalization after 75-percent epiboly. We performed some simulation trials suggesting that a continuous treadmill process would contribute to the motion of deep cells toward the vegetal pole. This hypothesis should be further tested and measurements from live experiments should be performed to definitely make a point. The involution/ingression of hypoblast could indeed be assessed, at least manually from digital specimens imaged in the appropriate orientation (such as our data ID 070927b, see Movie 6.1).

8.6.1 Hypotheses and Model

All the mechanisms involved in our model of gastrulation have been introduced in the previous case studies except for the margin contraction. In the “Intercalation Patterns” case study (Section 8.5), we presented the passive motion of the EVL-YSL margin when pushed by the marginal deep cells. We add here a new mechanism to model the “autonomous” sliding of the EVL-YSL margin starting at the time of hypoblast internalization and lasting until the tail bud closure, without contribution of the deep cells.

In the model, the margin yolk membrane particles simply transmit the “margin” label to their vegetal neighboring yolk membrane particles at a regular time interval T_{contract} , starting at 6 hours of the simulated time. This implementation is not a model of the mechanical margin contraction *per se*, but it allows to reproduce the observed progression of the margin.

8.6.2 Simulation, Parameter Space and Validation

All the scenarios that will be explored in the this case study have an initial state that is the direct output of the previous “Intercalation Pattern” case study, at 5.7 hours of the simulated time. Without any additional stage to our Waddingtonian Timeline, the epiboly motion would continue due to the radial intercalation of the deep cells. We give here a qualitative description of the mechanisms that might be

involved in the model but that are still in a phase of manual exploration at the time of writing this thesis.

We assume that the involution/ingression at the blastoderm margin to form the hypoblast requires two steps of differentiation. This is achieved by using a diffusive ligand secreted by the marginal yolk membrane particles. The deep cells interacting with an amount of ligand that exceeds the threshold differentiate i.e. ingress/invoke. The spreading of these cells on the yolk membrane involves a second differentiation step in response to a lower concentration of the same ligand, found in a more animal position. The hypoblast cells exert over their epiblast neighbors and the yolk a monopolar protrusive activity oriented toward the animal pole, leading to their antiparallel motion with respect to the epiblast cells.

The convergence-extension deformation is obtained by adding a polarization field originating from another ligand secreted in the dorsal region of the embryo. The epiblast cells perform a protrusive activity over their homotypic neighbors, oriented along the polarization field. The hypoblast cells also perform protrusive activity over their epiblast neighbors.

Measures

The following measures are proposed to test the different scenarios presented above:

- A measure of internalization: the temporal evolution of the ratio of the number of hypoblast cells over the number of epiblast cells.
- A measure of the animal progression of the hypoblast: the temporal evolution of the ratio between the length of the hypoblast over the animal-vegetal embryo height at the level of the bilateral symmetry plane on the ventral side (measure 1), and on the dorsal side (measure 2).
- A measure of the convergence: the temporal evolution of the dorso-ventral position of the deep cells could be done as proposed by Carmany-Rampey and Schier (Fig. 8.73).
- A measure of the pattern of singularities in the displacement vector fields: the work of Benoit Lombardot and David Pastor in our team provides the automated identification of singularities in the displacement vector field in both live and simulated data (see Fig. 7.6, Chapter 7). The fitness will be the volume of the intersection of the convex hull derived from the live and the simulated specimens.
- A measure of the cell flux along the animal-vegetal axis to detect the anti-parallel displacement of the epiblast and the hypoblast at the margin: this measure is expected to document the late epiboly treadmill issue.

- A measure of the spatio-temporal evolution of the cell division orientation: the orientation of cell division is given by the orientation of the dipole composed of the daughter cells' positions. The temporal evolution of the dipoles' orientations provides a relevant measure. The cell division orientation is defined in a referential centered on the mother cell, then on the center of the dipole. The referential axes are defined by the radius of the spherical approximation of the embryo (axis 1), the animal-vegetal axis (axis 2), and the orthogonal product of axis 1 and axis 2 (axis 3).

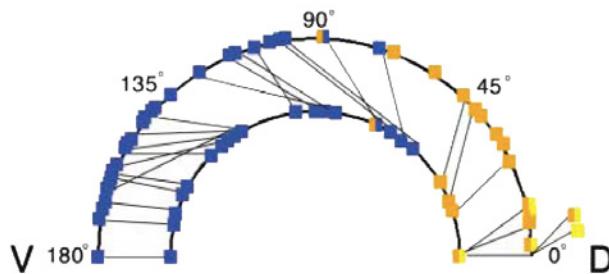


Figure 8.73: Dorso-ventral displacement of the cells during convergent extension. Each arc identifies individual cell positions along the dorso-ventral longitudes at a given time step. The black lines link the same individual cells at both time steps. The sum of the lengths of these links correlates with the speed of intercalation. Image and caption adapted from Carmany-Rampey and Schier [284]

8.6.3 Discussion

During the first 10 hours of development of the zebrafish, gastrulation is the most diversified episode in terms of morphogenetic deformation. The simultaneous action of epiboly, convergence, extension, and possibly internalization greatly complicates the analysis. For example, the closure of the tail bud can be due to EVL contraction, late radial intercalation, extension of the antero-posterior axis due to medio-lateral intercalation, or late epiblast-hypoblast treadmill. However, imposing constraints at the cellular level, such as protrusive activities, establishes a method of exploration of the causal mechanisms driving gastrulation in order to find the pattern of polarization field that could be responsible for the observed deformations.

CHAPTER 9

Conclusion

In this final chapter, we look back on the work accomplished so far and comment on its contribution to advances in developmental modeling, and the difficulties encountered. We discuss our design decisions, the validity and limitations of the results obtained—and the results not obtained but expected. In Section 9.1, we recapitulate the criteria for a “good” model. This is followed by a review of the choices that we made, both on the biomechanic side (Section 9.2) and the molecular/genetic side (Section 9.3). Finally, we globally reassess the case studies (Section 9.4) and conclude with directions for future work (Section 9.5).

9.1 Criteria for a “Good” Model

The most striking characteristic of multicellular systems is their deep *heterogeneity* of properties and behaviors. As Henri Atlan stated through his famous metaphorical phrase “Between Crystal and Smoke”[319], the structure of living matter is somewhere between the very regularly organized and the very randomly disorganized. The strongest principle in any attempt to model such systems should be to provide a framework that allows the capture this heterogeneity. We presented in Chapter 2 our view about the most suited level of organization for this task.

Microscopic models are too complex and uncontrollable

The molecular scale would be a tempting reductionist approach, since interactions between molecules are well described by models such as Molecular Dynamics. However, the hugeness of their number and the smallness of the timescale of their interactions render the practical simulation of multicellular tissue impossible with today’s computing power. The other drawback of such an extreme bottom-up approach would be the difficulty to interpret the generated dataset. Large-scale modeling and simulation of similar magnitude have been investigated in other fields of experimental science, such as the Blue Brain Project which now reproduces a cortical column containing tens of thousands of neurons and millions of synaptic contacts on highly detailed dendritic trees studded with ionic channels, and aims at scaling up to a whole human brain by the end of the next decade. There too, however, the twin problems of massive parameter tuning and “divergence of the emergence” remain as important as ever. In short: there are probably too many degrees of freedom and not enough *a priori* modeling assumptions.

Macroscopic models are too constrained and not expressive enough

On the other side of the spectrum, macroscopic formal descriptions offer great conciseness with general principles of conservation and continuity. Whether differential equations in continuous Euclidean space or statistical laws in state space, a few parameters can be sufficient to accurately reconstitute and predict certain dynamical systems phenomena. For example, the field of statistical mechanics has been successfully exported from thermodynamics and adapted to other well-mixed systems than gases such as sand dunes, bird flocks or human crowds. It can provide important analytical results about their dynamical regimes, typically “critical” parameter values at which “phase transitions” occur. In the case of developmental biology, disciplines of physics such as fluid or solid mechanics can provide interesting models of cell densities and collective trajectories, accounting for some aspects of tissue deformation. Here, however, the complexity and realism of these solutions is limited by the homogeneity of the parameters ruling the dynamics and the assumed continuity of the spatiotemporal distribution of variables. In contrast to the above scenario, we could say that the degrees of freedom are too few and the models too constraining.

Mesoscopic models provide the best compromise

This is why we believe that the best level to study morphogenesis in multicellular organisms is somewhere inbetween those two extremes. Continuous equations have to be coupled with additional local rules, and in the end necessitate a characterization of the diversity of the underlying cell types and behaviors. This approach requires a discrete representation of the tissue based on individual components. In our own study of early embryo development, we chose the single-cell level as it allows this local description of cell properties in relatively simple terms. The main criteria driving the choice of an appropriate dynamical framework are (a) the similarities between the variables involved in the model, on the one hand, and the observations and concepts construed and realized by the embryologists, on the other hand (discussed in Section 9.1.1); and (b) the simplicity of the description and the manageable number of parameters (discussed in Section 9.1.2). This should not only allow computationally feasible simulations by today’s standards but, most importantly, a biologically meaningful interpretation of their outcome.

9.1.1 Similarities Between Variables and Concepts

One could say that embryology itself has greatly “evolved” since the *Entwicklungsmechanik* revolution started by Wilhem Roux. He transformed a descriptive area into an experimental discipline, which explored the *causal* links of development—whether mechanical, chemical, thermal, or electrical—leading to developed organisms. The principal debate of that time revolved around whether parts self-differentiated or de-

pended on external influences. Since then, this puzzling question has been answered with increasing precision, first at the tissue level by the concept of “induction”, then at the cellular level with the advent of molecular biology. The great improvement of observation tools is now opening a new era that will master the automated integration of the cell fates of a complete organism (“*in toto*”). An experimental paradigm shift was triggered in 2006 with the mapping and reconstruction of all 959 cells of the *C. elegans*[293]. It now unfolds through projects and platforms such as BioEmergences in more complex species such as the zebrafish (see Section 7.1), to which we contributed our own methods of imaging data processing (see Section 7.2). While this is a work in progress, it offers a fundamentally new way of representing and understanding embryogenesis: the capacity of mapping a complete *cell lineage tree* augmented with local information such as cell positions [277], cell shapes, internal molecular states, gene expression through *cis*-regulatory reporters [320], and extracellular ligand concentrations. Based on these exhaustively reconstructed datasets, various correlations among quantitative observations, and among qualitative concepts, can then be detected and inferred by biologists.

This complex systems approach gives access to a new and vast empirical domain, which must be explored and charted from both the “classical” (see Section 1.1) and “theoretical” perspectives (see Section 2.1). Our point here is that the choice of a modeling paradigm must take into account this new phenomenology introduced by 3D+time reconstructions. Reconstructions and simulated embryos must “use the same vocabulary” to be able to define the grammar that links those two worlds. In practical terms, dynamical trajectories of the models gain in relevance if they reside in the same variable space as the live 3D+time reconstructions.

9.1.2 Computational Feasibility

The second criterion that drives the choice of a good modeling strategy is its suitability to current computer capacity. A balance has to be found between the number of variables and time steps, on the one hand, which impact the computational size and duration of a single embryo simulation, and the number of parameters that can be explored, on the other hand, which is combinatorially related to the total number of simulated embryos. It follows that these dimensions affect theoretical choices, too. For example, an *in toto* reconstruction of the zebrafish embryo until the bud stage, i.e. at about 25,000 cells and 10 hours of development, imposes much greater demands on the design of an encompassing model than a few dozen cells during a couple of hours. A concrete example can be found in the work of Sandersius, Newman et al. on the Subcellular Element Model (ScEM, see Section 3.1.1). The number of subcellular elements they use for single-cell studies is by far greater than in many-cell populations: typically 4,000 elements in a rheological experiment involving a single cell [224] versus only 20 elements per cell to model the large-scale vortices appearing during the primitive streak formation [184]).

In summary, these two criteria (variable-concept resemblance and practical computation) led us to make particular choices for (i) the cell biomechanics (discussed in Section 9.2) and (ii) the molecular/genetic regulation and signaling (discussed in Section 9.3).

9.2 Choices Concerning the Biomechanics

The biomechanical paradigm used in MECAGEN is a discrete description of cells represented by single particles. In Chapter 3, we presented our framework by distinguishing two steps in the elaboration of the model: first, the construction of the cell contacts (see Section 3.2.2), and second, the forces acting between neighboring cells (Sections 3.2.3-4). We discuss these choices here.

9.2.1 Spatial Neighborhood

Cell-cell interactions represented by particles must follow a strict principle: particles exert forces only over those particles with which they are in physical contact. No (direct) distant action similar to gravity or magnetic forces are applicable in the context of cell tissue—other than long-range axonal and dendritic synaptic transmission among neurons, but this dynamics is not yet in existence at the early embryonic stages. Propagation of mechanical constraints across a tissue can also be considered long-distance action, but it actually happens through local cell-to-cell contacts. In a model representation, this physical-contact principle translates into “*neighborhood links*” that must be defined and selected.

Metric neighborhoods are simple but can create unstable dynamics

A common and easy way to determine a spatial neighborhood is to set a maximal Cartesian distance within which cells can be in contact: this is the scope-limited “nearest neighbors” scheme, also called “metric” neighborhood, which we denoted by \mathcal{N}_i^m . Our early attempts to model cell-cell interactions were based on this metric notion, but we quickly observed that this approach had an important drawback: increasing the attraction part of the force potential (coefficient w_{adh}) inevitably led to unstable dynamics, where cells collapsed upon each other and eventually fused. The reason was simple: cells were simultaneously repelling their closest neighbors (those below the equilibrium distance r_{ij}^{eq}) and attracting the farther away neighbors (those beyond r_{ij}^{eq}). Problems happened when cells started attracting neighbors located *behind* other neighbors. In this situation, a large adhesion coefficient would trigger a “greedy” attraction of more and more neighbor cells, leading to a dramatic shrinkage and collapse of the tissue.

Topological neighborhoods are rational but computationally expensive

The solution we adopted to alleviate this problem was then to use a more realistic neighborhood definition that would take into account the nearby cells j truly “in contact” with a given cell i . This is the *topological* criterion \mathcal{N}_i^t , which formally defines the concept of being located “behind” something else. This solution has been demonstrated by the work of M. Meyer-Hermann and G. Schaller, who introduced the systematic use of the weighted Delaunay triangulation in a biomechanical simulation [213]. In 2D, the Delaunay triangulation algorithm allows to define neighborhood links based on a geometrical criterion, i.e. by forbidding the links that would produce triangles containing particles in their circumcircles. The 3D counterpart extends this criterion to the equivalent simplex: the tetrahedron. This approach, however, is computationally intensive as it has to deal with heavy data structures, so we decided to use instead a related algorithm based on the Gabriel graph [321]. This solution is easier to implement and faster to simulate, as the modified exclusion criterion is that every circle of which a neighborhood link is the diameter must not contain any other particles (than the link’s two extremities). We extended this criterion in 3D with spheres instead of circles. This strategy allowed us to avoid the collapsing behavior caused by high adhesion coefficients in the metric case.

9.2.2 Interaction Potential and Behavioral Forces

Once spatial relationships between neighbors are established, the next step is to define “*forces*” responsible for the mechanical interactions. The main principle of the motion equation is that, in multicellular systems, the cells are so small and their interactions so sticky that their motion is not well described at their scale by classical mechanics. Instead of the acceleration, it is rather the velocity of an element that is proportional to the sum of the forces applied to it. The consequence is that mass-based inertia plays no part in cell motion, which is rather due to the instantaneous action of forces [231]. This principle is widely adopted in various models of multicellular assembly [219][183][221]. For the definition of the forces *per se*, we decided to dissociate two types of interactions: “passive” adhesion/repulsion interactions responsible for the relaxed state of the assembly, and “active” forces (such as protrusion) responsible for the individual behavior of the cells.

“Passive” adhesion/repulsion interaction potential

Concerning the first type of interactions, we explored several models of potentials offering long-range attraction and short-range repulsion: the Morse potential [183], the Lennard-Jones potential [322], Dissipative Particle Dynamics [323], and contact mechanic models derived from the Hertz model such as Johnson-Kendall-Roberts [324]. The conclusion that we drew from this practical study was that, even if each model had its own specificity as it was designed to study other phenomena, they

all shared the same peculiarity and limitation the single-particle cell framework: here, the repulsion part of the potential is responsible for both the integrity of the cellular domain and the stiffness of the response. Thus it is difficult to pretend that the interaction potential model correctly integrates the biomechanical properties of cortical tension in a dense multicellular context. Other single-particle cell models have added a *volume conservation* force to avoid cells becoming smaller when the adhesion coefficient increases, even in the topological neighborhood context [213]. In our study, we decided to select a simple solution, in which elastic forces are scaled by our estimation of the surface of contact. In addition to displaying a similar profile (Fig. 3.30) with a vanishing force beyond a maximum distance limit, this solution also properly deals with rescaling the force when cells' size is dramatically reduced during the early stages of development. The repulsion coefficient w_{rep} controls both the cell's volume conservation and stiffness of response until equilibrium, as these two properties are inseparable. The adhesion coefficient w_{adh} controls the adhesive property of the cells beyond equilibrium.

“Active” behavioral forces

The second type of interaction forces introduced in this work models the specialized proactive behaviors of the cells. Within the possible repertoire of a cell behavior ontology (CBO, Section 5.1.1.), we focused in this study on a central mechanism that we extensively used in the cases studies in Chapter 8: mesenchymal cell *protrusion*. Protrusive behavior constitutes a principle that is rarely used in multicellular studies (one recent publication, posterior to and independent from us, uses it for the same purpose [216]). Yet, it plays a key role in active migration and intercalation [325][326][327]. Several models have proposed mechanisms to explain this behavior in a single cell, for example Weliky et al. introduced an explicit deformation of the cell shape in their study of the convergence-extension in the *Xenopus* notochord [182]. We believe, however, that a formalization of this behavior makes sense only if the protruding cell has a support to exert its action upon. It means that models of cell migration should always associate to the migration force intrinsically generated by a protruding cell a counterpart force on its physical support (i.e. the surrounding cells in the neighborhood) oriented in the opposite direction. Without this counterpart force, the model would be like the action of an external actor pushing the cell in their migratory direction. This principle, when applied to an assembly of protruding cells, creates a non-trivial emergent behavior of tissue deformation and legitimates in itself the need of computer simulation to observe and analyze it. We defend the notion that this mechanism could explain various embryonic phenomena observed during the zebrafish early development (Case Studies 8.5, 8.6). Another interesting approach is proposed by Sandersius et al. in their Subcellular Element Model framework (ScEM) [184]: they model the process of cytoskeletal polymerization involved in protrusion by adding subcellular elements at the tip of the protruding pole

and removing other subcellular elements at the other end of the cell. This mechanism, coupled with cell-cell adhesion, also induces a migratory behavior in the cell population.

9.2.3 External Objects

One of the main advantages of particle-based modeling (whether single- or multiple-particle) is the possibility to seamlessly aggregate to the multicellular model external physical objects that have their own dynamical properties. This point is illustrated in our thesis with the addition of the large yolk cell supporting zebrafish development. It is clear that no relevant insight on the mechanics of the zebrafish development could be gained without taking into account the yolk structure. The particle physics framework can alleviate the discontinuities generated by these external structures relatively easily: they, too, are modeled with particles (here, the yolk cell is a high-degree polyhedron) and contacts between structure particles and cells are exactly of the same kind as cell-cell contacts, only with different force parameters. One could compare this framework to “Lego” bricks, which despite their different shapes and sizes all have the same pegs and holes by which they can attach to each other. In future studies, this standardized framework can be extended to other types of extracellular interactions in a straightforward manner.

9.2.4 Single-Cell vs. Subcellular Mechanics

The single-particle cell model also has drawbacks, however. Among them, the “coarse-graining” performed at this level of abstraction is too high to integrate the subcellular structures involved in the mechanics of embryonic morphogenesis. An example of this is the extra-cellular matrix (ECM) that has been shown to participate in the establishment of morphogenetic domain boundaries [328]. ECM is a dynamical network of filament molecules such as fibronectin and fibrillin [329] secreted by the cells. Despite this, the absence of ECM in the model can be partly justified by the results of a recent study showing that it displays actually little relative motion with respect to the surrounding cells, both ECM and cells moving together as a single mass [330]. This would imply, conversely, that cells apply only little force on the ECM because, similarly to our cell-cell protrusion mechanism, any force would induce ECM motion in the opposite direction, which is not observed. The mechanical influence of the ECM also appears limited in the early stages of zebrafish development as its expression is only observed after 80 percent epiboly, coinciding with the convergence-extension episode [331].

9.2.5 Spherical-Voronoi Approximation

Another issue with the single-particle paradigm is the lack of explicit cell shape description, making it necessary to use an approximation (i.e. a statistical average)

to calculate the contact surface area between cells. In the project, as we lacked real embryo membrane segmentations during the later stage of zebrafish development, we opted for a empirical law of the surface defined by weighted Voronoi domains based on abstract numerical experiments. We expect that this approximation is acceptable when cells are densely packed [332], but when they present “exposed” facets (i.e. not touching other cells), the hybrid spherical-voronoi approximation of the particles does not properly account for cell shapes. This is particularly striking during the first stages of the zebrafish development, where blastomeres form large cells spread out on the surface of the yolk. There, the model would benefit from finer-grained representations such as the ones provided by the Subcellular Element Model or the Cellular Potts Model.

9.3 Choices Concerning the Molecular/Genetic Regulation and Signaling

This part of our model follows the same criteria of vocabulary and computational suitability. The interface we use here to specify the parameters is the BioTapestry software, which is the most popular visualization tool for gene regulatory networks (GRNs) involved in animal development. We believe that this strategy should facilitate the communication with researchers familiar with this environment. From the theoretical perspective, in order to describe the regulation kinetics, we preferred the conventional system of ordinary differential equations (ODEs) over other paradigms. For sure, more accurate simulations of the molecular dynamics could be obtained from finer-grained frameworks such as stochastic event-based models using Gillespie’s method [333] or rule-based models like the Kappa language representing protein interaction networks. It has been shown that important molecular dynamical phenomena occur when the number of molecules is low [334], as even individual molecules can play a role in the stability of the cell state. Thus models at the molecular scale might better capture this singular dynamics than ODEs describing average concentrations. Yet again, just like in the mechanical part, their computational cost is also higher and does not suit our objectives in terms of simulation time. Looking in the opposite direction, even faster simulations could be obtained with Boolean networks instead of real-value variables, with adequate performance fitness defined on real embryo datasets [239]. But while Boolean network are well suited to the *cis*-regulatory logic of GRNs (which we included in our model, see Chapter 4), concerning other components of molecular and genetic regulation, such as protein-protein interactions or cell signaling, the parametric space of ODEs is better adapted to direct experimental measures and validation.

9.4 Assessment of the Case Studies

Translating state-of-the-art biological knowledge into models

The life sciences have generated a remarkable wealth of knowledge on multiple scales. A model such as the one that we presented in this thesis exploits only a small fraction of this immense corpus. Many improvements are still possible within the currently available “state of the art”. Yet, modeling also implies *selecting* specific subparts of this information, and requires an extensive production of experimental measures to calibrate the various parameters involved in the equations. Thus, on the modeling side, the main problem is the high number of parameters, making their exploration time consuming and difficult.

The strategy that we adopted in the case studies was to separate the parameters into two categories in each case study: ones set to fixed values based on qualitative estimations, and others left open to exploration. Naturally, the choices presiding over the first category are subjective and debatable. It is possible that changing the values of some of the “fixed” parameters, i.e. transferring them to the second category, could generate simulated phenotypes that would perform equivalently or even better than the ones subject to reevaluation.

On the positive side, this subjective, “hand-made” aspect confers a particular meaning to the important mechanisms that are involved. On the negative side, it also signals a lack of systematicity or exhaustiveness which would be ideally required for an in-depth understanding of the model and, through the model, the biological mechanisms themselves. A classical solution is to calibrate some of the parameters via direct measurements obtained from the real systems. In our project, this calibration phase was not done because (a) the direct relationship between most of the parameters and their real counterparts needed further analysis, and (b) some of the experimental measures were just not yet available.

- (a) The calibration phase of the mechanical parameters of the model could be done through specific *in silico* experiments aiming at characterizing the microscopic behavior of the system. We may cite as an example the values attributed to the adhesion and repulsion coefficient of the passive relaxation forces. In the project, this evaluation was reverse-engineered via an evaluation of the macroscopic behavior of the system. In the first case study, “Investing the Yolk Biomechanical Properties” (Section 8.1), the optimal values were obtained through a macroscopic perturbation of the system—in this case, a pressure exerted over the yolk surface. This type of the study could be generalized at the cell level through systematic perturbation experiments applied on single cells. This would be similar to the adhesion measures in germ-layer progenitors by single-cell force spectroscopy in Krieg et al. [335], or at the tissue level, to the study of the effective viscosity of cell aggregates in Marmottant et al. [336].

- (b) Most parameters used in the model would also require *in vivo* measurements that have not been realized yet. Two of the most important parameters in MECAGEN are the intensity f^A of the protrusive forces and the polarization axes \vec{U}_i . Calibrating the protrusive forces could already be feasible with *in vitro* measures of migrating individual cells. Experiments have been designed to quantitatively assess the cellular traction stresses by traction force microscopy in 2D culture cells [337]. It has also been done in a 3D environment as shown in a recent study of 3T3 fibroblasts migration [338]. This study aims at deciphering the protrusive force patterns occurring in cell-ECM migration. The next step would be to apply this methodology in a multicellular assembly, first *in vitro* on cells extracted from developing embryos, then eventually *in vivo*. Calibrating the polarization axes appears to be a more immediate challenge, as the conjunction of intracellular molecular markers of polarization and large-scale *in vivo* reconstruction platforms like BioEmergences should produce 3D+time maps of the polarization fields at the scale of the embryo.

What is needed from biology to make a modeling endeavor successful

The whole process of calibrating the parameters would be facilitated by provided access to databases of reconstructed embryos. The sophisticated reconstruction methods that are being developed today to generate *in toto* measurements of the spatial-temporal properties of the cells in developing embryos will soon become standard tools in the lab of tomorrow. Today's solutions, however, are not yet at this stage of completion or generality, as they are still very much dependent on particular microscope setups. Future advances in microscope technology, both in the acquisition time and resolution dimensions, should produce images that could be reconstructed on plain laptops. In the meantime, theoretical studies of development would greatly benefit from an easier access to the existing reconstructed embryos. This request raises two challenges: the computational infrastructure and, more importantly, the standardization of the storage format. This latter point meets the challenge of designing a “Cell Behavior Ontology” that would be accepted by the whole community.

The persistent enigmas of zebrafish early development

Among the multiple questions that the first ten hours of zebrafish development raise, one of the most puzzling concerns the shape transition observed at the sphere stage. Although it has been observed and recognized for decades now, this famous transition characterized by a flattening of the yolk-blastoderm interface has not been studied extensively enough by the zebrafish research community. In the fifth case study, about “Intercalation Patterns” (Section 8.5), we were simply unable to generate a simulated phenotype that would resemble the sphere stage embryo. This may be due to the existence of an additional, unknown mechanism that was not

taken into account in our model. To our knowledge, however, no extra mechanism has been proposed or hypothesized by developmental biologists. Since this transition does not seem to produce a new apparent embryonic structure, this deformation could be the consequence of an unknown mechanism by which the zebrafish would “skip” the sphere stage (simply transitioning from the oblong stage to dome stage) without losing a morphological structure. However, this transition appears to be conserved in teleost development, as illustrated by the developmental stage 11 shown in [339].

The help of agent-based modeling in elucidating those questions

The sphere stage transition is the only developmental stage that poses a problem in the first 10 hours of the zebrafish development. All the other major morphogenetic events are likely to be reproduced by the implemented rules presented in this project. This highlights the capacity of the MECAGEN simulation platform to handle the great number of heterogeneous interactions occurring during embryonic development. However, the current state of this study is situated somewhere between a “proof of concept” validating the methodology and the actual realization of the complete morphogenetic development that is planned for the future.

About reconstructed data and measures that were not exploited by our model

One type of measures that was evoked in Section 1.3 about methodological considerations but not been exploited in the case studies comprise the “extensive microscopic” measures. They consist of *in toto*“fields”, ranging from local displacement vectors to more structured objects such as rearrangement pattern tensors [340]. This type of measure raises the issue of the automated spatiotemporal registration of embryos (simulated or live). In the third case study, “Shaping the Blastula” (Section 8.3), we mentioned that we planned on superimposing the spatiotemporal coordinates of the cell centers from simulated and live specimens (such as shown in Movie 8.7). This registration could be done manually by visual estimation of the geometrical transformation parameters. However, the ideal solution necessitates the design of a *prototype embryo*. This concept is based on the idea that all developing embryos are mere variations on a single structural theme, the “prototype”, whose components may slightly vary in each individual specimen. Designing the prototypical embryo is an important project that has not been realized yet, and would itself require building macroscopic measures. Using data-mining methods to “learn” the different associations between these observations would allow to automatically recognize the morphogenetic events occurring in a phenotypic trajectory. The spatiotemporal coordinates of the embryonic structures would thus be associated to the other macroscopic measures. Such a system could then detect the macroscopic features extracted from raw data imaging (e.g. an internalizing tissue in the dorsal region) and

recover their spatiotemporal coordinates in a prototypic template (e.g. the imaged data have the spatial coordinates of the generic internalization shield and the associated timing is 6 hpf). This prototyping tool would contribute to great advances in the comparison of embryos and the fusion of partial heterogeneous datasets.

Installing and exploiting an evolutionary search

The final parameter exploration carried out in our various case studies was performed by regularly sampling or “sweeping” parameter space, and producing more or less complete landscapes of the associated measures or fitness functions. This strategy is far from optimal if we aim at finding the “best” set of parameters, especially in higher dimensions and larger intervals. Other strategies of optimization, such as gradient descent, simulated annealing, or probably best if all, an *evolutionary search*, would be more efficient as they would converge faster toward a solution. However, as we were not only trying to find the best set of parameter, but also understand the relationship between the parameters and macroscopic measures, we decided to survey and chart whole regions of parameter space—which was difficult in 4D and beyond. Advanced “genetic algorithms” evaluating multi-objective functions could prove a fruitful direction to improve our treatment of parameter space, especially in higher dimensions. But the problem of visualization would remain. Another issue with automated exploration techniques is that they may exploit every flaw or idiosyncratic aspect of the measures in the “wrong” way, meaning that they can find hyperspecialized solutions that perform really well but cannot be interpreted or generalized (e.g. the margin-rolling phenotype observed in Case Study 8.5, see Movie S.7).

9.5 Future Work

9.5.1 Portability to Other Animal Models

The MECAGEN project was envisioned as a generic modeling platform for all types of animal development. Its foundational principles should be applicable to any multicellular system involving biomechanics and molecular/genetic regulation and signaling. The next step in the development of this platform is to expand its cell behavior ontology (CBO) to integrate other types of epithelial behaviors. Not many improvements should be necessary to fulfill this objective. The adhesion coefficient of “lateral” neighborhood links could be increased and the epithelial behaviors mentioned in Section 5.1.1., apical constriction and active intercalation in the epithelial sheet, should allow the platform to simulate monolayered embryos such as *Drosophila*. Apical constriction can be modeled by adding couples of intrinsic and extrinsic behavioral forces, $\vec{F}^{A,int}$ and $\vec{F}^{A,ext}$, where the latter must be oriented toward the apico-basal axis above its apical region. This simple mechanism should

trigger the invagination of the epithelium. Cell intercalation in the *Drosophila* epithelium is thought to operate by local junction remodeling [341] and involve a molecular mechanism that differs from the protrusive behavior observed in mesenchymal cells. Yet, we believe that this mechanism could also be modeled by the same couple of behavioral forces that we used in mesenchymal protrusion. The only difference would consist of imposing that the polarization axis supporting this intercalation be constrained in the epithelial plane.

This being said, however, the current MECAGEN implementation is also not compatible with every animal cell behavior. Its major limitation resides in the single-particle cell abstraction, which is not adapted to cellular shapes that depart greatly from spheroids or cuboids. Other types of animal morphogenesis are based on cell elongation, as in *Phallusia* or nematodes. In those cases, an ellipsoid particle model such as Palsson's [211] provides a reasonable solution for asymmetric cells. In later stages of vertebrates, too, cells differentiate into extremely stretched shapes such as muscle cells, for which the single-particle framework is clearly inappropriate. As mentioned previously, another solution is to descend to a lower scale involving a finer-grain description based on subcellular elements, which should be able to mimic those unconventional cell shapes.

9.5.2 Interactive Exploration of Parameter Space

Exploring parameter space in the case studies was not without frustration, as the interpretation of the fitness landscapes did not clearly point to the underlying phenotypes associated to each measured value. A future avenue of research that should facilitate this exploration would include the design of *interactive* landscapes. One can envision fitness landscapes that would be “clickable” by a human user, and the selection of certain points in that space would display the associated phenotypes in the measure space (e.g. the temporal plot of the doming macroscopic measures shown in Section 8.5), and also in the variable space via 3D rendering of the developmental trajectories. This work would require much more software engineering but all the elements are in place to rapidly develop a workable solution. The phenotypic visualization could be pre-calculated but in the case of thousands of embryos that each comprises millions of time-dependent variables, new simulations would be triggered by this hand-selection of parameters. To our knowledge, no interactive exploration solution of this type exists yet, although it would be of great benefit to the study of all kinds of complex dynamical systems.

9.5.3 Toward an “Evo-Devo” Perspective

Another promising extension of this work, once the “MECA” and “GEN” parts of the model are finally connected and their dynamics truly coupled, will be its potential application to “evolutionary developmental” questions (mentioned in the introduction of Section 2.2). Evo-devo does not generally aim at directly comparing

simulated phenotypes and real embryos but rather asking how evolution is able to generate new structures and behaviors at the level of the whole embryo. This would involve modeling and simulating a higher level of organization yet, in which the whole embryo behaves and potentially interacts with its environment. It can be seen as another way to explore the model's parameter space without isolating the developmental trajectories around a target phenotype (zebrafish in this project). The evaluation would emerge from a Darwinian selection process in an artificial environment—whether by survival or reproduction of particular behaviors. A particularly fascinating exploration would be to start from a simple individual organism such as the *urbilaterian*, a hypothetical ancestor of all animals exhibiting bilateral symmetry [342][343], and guide its evolution through mutation of its gene regulatory networks toward multiple descendant species. To our knowledge, no computational evo-devo study has proposed a developmental model presenting as many similarities with real biological systems as MECAGEN. It would be a great opportunity to better decipher and understand the evolutionary forces that drive the diversity of life.

Bibliography

- [1] Enrico Coen. *The Art of Genes*. How Organisms Make Themselves. Oxford University Press on Demand, 1999. (Cited on pages 13 and 53.)
- [2] Nick Hopwood. Embryology. *The Cambridge History of Science. Volume 6*, 2008. (Cited on pages 14, 15 and 16.)
- [3] Joseph Needham. *Chemical Embryology*. CUP Archive, 1931. (Cited on page 14.)
- [4] Aristotle. *Aristotle - Works*. Oxford Clarendon Press, November 1908. (Cited on page 14.)
- [5] T Horder. History of Developmental Biology. *Encyclopedia of Life Sciences (ELS)*. John Wiley & Sons, Ltd: Chichester, 2010. (Cited on page 15.)
- [6] Rene Descartes. L'Homme de René Descartes et la formation du foetus, avec les Remarques de L. de La Forge, à quoy l'on a ajouté Le Monde ou traité de la lumière du mesme auteur [Publ. par Clerselier et suivi de la trad. de la préf. de Schuy]., 1677. (Cited on page 15.)
- [7] Jacques Roger. *The Life Sciences in Eighteenth-Century French Thought*. Stanford University Press, March 1998. (Cited on page 15.)
- [8] M De Felici and G Siracusa. The rise of embryology in Italy: from the Renaissance to the early 20th century, 2000. (Cited on page 15.)
- [9] Robert Hooke. *Micrographia - 1665*. Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses with Observations and Inquiries Thereupon. Project Gutenberg, March 2005. (Cited on page 15.)
- [10] H Gest. The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society. *Notes and Records of the Royal Society*, 58(2):187–201, May 2004. (Cited on page 15.)
- [11] David Lagunoff. Portraits of science. A Polish, Jewish scientist in 19th-century Prussia. *Science*, 298(5602):2331, December 2002. (Cited on page 16.)
- [12] Robert Remak. *Untersuchungen über die Entwicklung der Wirbelthiere*. Reimer, Berlin, November 1855. (Cited on page 16.)
- [13] Albert Kölliker. *Entwickelungsgeschichte Des Menschen Und Der Höheren Thiere*. Verlag Von Wilhelm Engelmann, September 1861. (Cited on page 17.)
- [14] Nick Hopwood. Approaches and species in the history of vertebrate embryology. *Methods in molecular biology (Clifton, NJ)*, 770:1–20, 2011. (Cited on page 17.)
- [15] V Hamburger. *Wilhelm Roux: visionary with a blind spot.*, volume 30. Kluwer Academic Publishers, 1997. (Cited on page 17.)
- [16] Peter A Lawrence and Michael Levine. Mosaic and regulative development: two faces of one coin. *Current Biology*, 16(7):R236–R239, April 2006. (Cited on pages 17 and 18.)
- [17] E H Davidson. Spatial mechanisms of gene regulation in metazoan embryos. *Development*, 113(1):1–26, September 1991. (Cited on page 17.)
- [18] R G Harrison. Experiments on the development of the fore limb of Amblystoma, a self differentiating equipotential system. *Journal of Experimental Zoology*, 1918. (Cited on page 18.)
- [19] E M De Robertis, E A Morita, and K W Cho. Gradient fields and homeobox genes. *Development*, 112(3):669–678, July 1991. (Cited on page 18.)
- [20] J Huxley and G De Beer. *The elements of experimental embryology*. Cambridge comparative physiology. The Cambridge University Press, 1934. (Cited on page 18.)

- [21] H Spemann and Hilde Mangold. über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. *Wilhelm Roux' Archiv für Entwicklungsmechanik der Organismen*, 100(3-4):599–638, September 1924. (Cited on page 18.)
- [22] J Holtfreter. Reminiscences on the life and work of Johannes Holtfreter. *Developmental biology (New York, N.Y. : 1985)*, 7:109–127, 1991. (Cited on pages 18 and 19.)
- [23] Conrad Hal Waddington. *Organisers & genes*. The University Press, 1940. (Cited on page 18.)
- [24] S F Gilbert. Induction and the Origins of Developmental Genetics. In *Developmental Biology*. <http://9e.devbio.com>, 1991. (Cited on pages 18 and 20.)
- [25] Rony Armon. Between Biochemists and Embryologists The Biochemical Study of Embryonic Induction in the 1930s. *Journal of the History of Biology*, 45(1):65–108, December 2010. (Cited on page 19.)
- [26] J Gerhart. Johannes Holtfreter: January 9, 1901–November 13, 1992. *National Academy of Sciences (US)*, 73:209–228, 1998. (Cited on page 19.)
- [27] H Steinbeisser. The impact of Spemann's concepts on molecular embryology. *The International journal of developmental biology*, 40(1):63–68, February 1996. (Cited on page 19.)
- [28] J Needham, C H Waddington, and D M Needham. Physico-chemical experiments on the amphibian organizer. *Proceedings of the Royal Society of London Series B*, 114(789):393–422, 1934. (Cited on page 19.)
- [29] E Wehmeier. Versuche zur Analyse der Induktionsmittel bei der Medullarplatteninduktion von Urodelen. *Development genes and evolution*, 132(2):384–423, 1934. (Cited on page 19.)
- [30] L G Barth and S Graff. The chemical nature of the amphibian organizer. *Cold Spring Harbor Symposia on Quantitative Biology*, 6(0):385–391, January 1938. (Cited on page 19.)
- [31] D S Kessler and D A Melton. Vertebrate embryonic induction: mesodermal and neural patterning. *Science*, 266(5185):596–604, October 1994. (Cited on page 19.)
- [32] Lauri Saxen. Transfilter neural induction of amphibian ectoderm. *Developmental Biology*, 3:140–152, April 1961. (Cited on page 19.)
- [33] H Tiedemann, F Lottspeich, M Davids, S Knöchel, P Hoppe, and H Tiedemann. The vegetalizing factor. A member of the evolutionarily highly conserved activin family. *FEBS letters*, 300(2):123–126, March 1992. (Cited on page 19.)
- [34] Thomas Hunt Morgan. *The Theory of the Gene*. Yale University Press, September 1926. (Cited on page 19.)
- [35] S Gluecksohn-Schoenheimer. The Development of Two Tailless Mutants in the House Mouse. *Genetics*, 23(6):573–584, November 1938. (Cited on page 19.)
- [36] S Gluecksohn-Schoenheimer. The Effect of an Early Lethal (t) in the House Mouse. *Genetics*, 25(4):391–400, July 1940. (Cited on page 19.)
- [37] C Nüsslein-Volhard and E Wieschaus. Mutations affecting segment number and polarity in *Drosophila*. *Nature*, 287(5785):795–801, October 1980. (Cited on pages 19, 22 and 52.)
- [38] C Nüsslein-Volhard. The zebrafish issue of Development. *Development*, 139(22):4099–4103, October 2012. (Cited on page 19.)
- [39] Jonathan M W Slack. Conrad Hal Waddington: the last Renaissance biologist? *Nature Reviews Genetics*, 3(11):889–895, November 2002. (Cited on page 20.)
- [40] François Jacob and Jacques Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology*, 3:318–356, June 1961. (Cited on page 20.)
- [41] C H Waddington. *Principles of Embryology*. The Macmillan Company, 1956. (Cited on page 21.)

- [42] M I Arnone and E H Davidson. The hardwiring of development: organization and function of genomic regulatory systems. *Development*, 124(10):1851–1864, May 1997. (Cited on page 21.)
- [43] Eric H Davidson. *The Regulatory Genome*. Gene Regulatory Networks in Development And Evolution. Academic Press, 2006. (Cited on page 21.)
- [44] Tzu-Min Chan, William Longabaugh, Hamid Bolouri, Hua-Ling Chen, Wen-Fang Tseng, Chung-Hao Chao, Te-Hsuan Jang, Yu-I Lin, Shao-Chin Hung, Horng-Dar Wang, and Chiou-Hwa Yuh. Developmental gene regulatory networks in the zebrafish embryo. *Biochimica et biophysica acta*, 1789(4):279–298, April 2009. (Cited on pages 21, 114 and 151.)
- [45] F Pelegri. Maternal factors in zebrafish development. *Developmental dynamics : an official publication of the American Association of Anatomists*, 228(3):535–554, 2003. (Cited on page 21.)
- [46] J P Renard, P Baldacci, V Richoux-Duranthon, S Pournin, and C Babinet. A maternal factor affecting mouse blastocyst formation. *Development*, 120(4):797–802, April 1994. (Cited on page 21.)
- [47] E R Gavis and R Lehmann. Localization of nanos RNA controls embryonic polarity. *Cell*, 71(2):301–313, October 1992. (Cited on page 21.)
- [48] D St Johnston and C Nüsslein-Volhard. The origin of pattern and polarity in the Drosophila embryo. *Cell*, 68(2):201–219, January 1992. (Cited on page 21.)
- [49] B Bowerman, B A Eaton, and J R Priess. skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early *C. elegans* embryo. *Cell*, 68(6):1061–1075, March 1992. (Cited on page 21.)
- [50] Aniket V Gore, Shingo Maegawa, Albert Cheong, Patrick C Gilligan, Eric S Weinberg, and Karuna Sampath. The zebrafish dorsal axis is apparent at the four-cell stage. *Nature*, 438(7070):1030–1035, December 2005. (Cited on pages 21 and 165.)
- [51] André Pires-daSilva and Ralf J Sommer. The evolution of signalling pathways in animal development. *Nature Reviews Genetics*, 4(1):39–49, January 2003. (Cited on page 22.)
- [52] R Nusse and H E Varmus. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell*, 31(1):99–109, November 1982. (Cited on page 22.)
- [53] R P Sharma. Wingless a new mutant in *Drosophila melanogaster*. *Drosophila information service*, 1973. (Cited on page 22.)
- [54] S Artavanis-Tsakonas, M D Rand, and R J Lake. Notch signaling: cell fate control and signal integration in development. *Science*, 284(5415):770–776, April 1999. (Cited on page 22.)
- [55] Sarah J Bray. Notch signalling: a simple pathway becomes complex. *Nature reviews Molecular cell biology*, 7(9):678–689, September 2006. (Cited on page 22.)
- [56] Bernhard Fuss and Michael Hoch. Notch signaling controls cell fate specification along the dorsoventral axis of the Drosophila gut. *Current biology : CB*, 12(3):171–179, February 2002. (Cited on page 22.)
- [57] C Wicking, I Smyth, and A Bale. The hedgehog signalling pathway in tumorigenesis and development. *Oncogene*, 18(55):7844–7851, December 1999. (Cited on page 22.)
- [58] Jin Jiang and Chi-Chung Hui. Hedgehog signaling in development and cancer. *Developmental cell*, 15(6):801–812, December 2008. (Cited on page 22.)
- [59] Mary Y Wu and Caroline S Hill. Tgf-beta superfamily signaling in embryonic development and homeostasis. *Developmental cell*, 16(3):329–343, March 2009. (Cited on page 22.)
- [60] Young-Kyung Bae, Nathanie Trisnadi, Snehalata Kadam, and Angelike Stathopoulos. The role of FGF signaling in guiding coordinate movement of cell groups: guidance cue and cell adhesion regulator? *Cell Adhesion & Migration*, 6(5):397–403, September 2012. (Cited on page 22.)

- [61] A Wayne Orr, Brian P Helmke, Brett R Blackman, and Martin A Schwartz. Mechanisms of mechanotransduction. *Developmental cell*, 10(1):11–20, January 2006. (Cited on page 22.)
- [62] Diana E Jaalouk and Jan Lammerding. Mechanotransduction gone awry. *Nature reviews Molecular cell biology*, 10(1):63–73, January 2009. (Cited on page 22.)
- [63] Michele A Wozniak and Christopher S Chen. Mechanotransduction in development: a growing role for contractility. *Nature reviews Molecular cell biology*, 10(1):34–43, January 2009. (Cited on page 22.)
- [64] Jeroen Eyckmans, Thomas Boudou, Xiang Yu, and Christopher S Chen. A Hitchhiker’s Guide to Mechanobiology. *Developmental cell*, 21(1):35–47, July 2011. (Cited on page 22.)
- [65] J Wolff. *Das Gesetz der Transformation der Knochen*. Hirschwald, 1892. (Cited on page 22.)
- [66] Wilhelm Roux. *Gesammelte Abhandlungen über Entwicklungsmechanik der Organismen*. W. Engelmann, 1895. (Cited on page 22.)
- [67] D A W Thompson. *On Growth and Form*. Cambridge University Press, 1917. (Cited on page 22.)
- [68] R P Franke, M Gräfe, H Schnittler, D Seiffge, C Mittermayer, and D Drenckhahn. Induction of human vascular endothelial stress fibres by fluid shear stress. *Nature*, 307(5952):648–649, February 1984. (Cited on page 22.)
- [69] Celeste M Nelson, Ronald P Jean, John L Tan, Wendy F Liu, Nathan J Sniadecki, Alexander A Spector, and Christopher S Chen. Emergent patterns of growth controlled by multicellular form and mechanics. *Proceedings of the National Academy of Sciences of the United States of America*, 102(33):11594–11599, August 2005. (Cited on page 22.)
- [70] Nikolce Gjorevski and Celeste M Nelson. Endogenous patterns of mechanical stress are required for branching morphogenesis. *Integrative Biology*, 2(9):424–434, September 2010. (Cited on page 22.)
- [71] Adam C Martin, Michael Gelbart, Rodrigo Fernandez-Gonzalez, Matthias Kaschube, and Eric F Wieschaus. Integration of contractile forces during tissue invagination. *The Journal of cell biology*, 188(5):735–749, March 2010. (Cited on page 22.)
- [72] Rodrigo Fernandez-Gonzalez, Sérgio de Matos Simoes, Jens-Christian Röper, Suzanne Eaton, and Jennifer A Zallen. Myosin II dynamics are regulated by tension in intercalating cells. *Developmental cell*, 17(5):736–743, November 2009. (Cited on pages 22 and 39.)
- [73] Gregory F Weber, Maureen A Bjerke, and Douglas W Desimone. A mechanoresponsive cadherin-keratin complex directs polarized protrusive behavior and collective cell migration. *Developmental cell*, 22(1):104–115, January 2012. (Cited on pages 22 and 131.)
- [74] Bernhard Hampelz, Yannick Azou-Gros, Roxane Fabre, Olga Markova, Pierre-Henri Puech, and Thomas Lecuit. Microtubule-induced nuclear envelope fluctuations control chromatin dynamics in Drosophila embryos. *Development*, 138(16):3377–3386, August 2011. (Cited on page 22.)
- [75] Nicolas Desprat, Willy Supatto, Philippe-Alexandre Pouille, Emmanuel Beaurepaire, and Emmanuel Farge. Tissue deformation modulates twist expression to determine anterior midgut differentiation in Drosophila embryos. *Developmental cell*, 15(3):470–477, September 2008. (Cited on pages 22, 39 and 128.)
- [76] Leif C Lindeman, Cecilia L Winata, Hvard Aanes, Sinnakaruppan Mathavan, Peter Alstrom, and Philippe Collas. Chromatin states of developmentally-regulated genes revealed by DNA and histone methylation patterns in zebrafish embryos. *The International journal of developmental biology*, 54(5):803–813, 2010. (Cited on page 23.)

- [77] Leif C Lindeman, Ingrid S Andersen, Andrew H Reiner, Nan Li, Håvard Aanes, Olga Østrup, Cecilia Winata, Sinnakaruppan Mathavan, Ferenc Müller, Peter Aleström, and Philippe Collas. Prepatternning of Developmental Gene Expression by Modified Histones before Zygotic Genome Activation. *Developmental cell*, pages 1–12, November 2011. (Cited on page 23.)
- [78] Daniel H Chitwood and Marja C P Timmermans. Small RNAs are on the move. *Nature*, 467(7314):415–419, September 2010. (Cited on page 23.)
- [79] R Keller. Physical Biology Returns to Morphogenesis. *Science*, 338(6104):201–203, October 2012. (Cited on page 23.)
- [80] Johannes Holtfreter. *Gewebeaffinität, ein Mittel der embryonalen Formbildung*. Unfound publisher, 1939. (Cited on page 23.)
- [81] M S Steinberg. On the mechanism of tissue reconstruction by dissociated cells. I. Population kinetics, differential adhesiveness, and the absence of directed migration. *Proceedings of the National Academy of Sciences of the United States of America*, 48:1577–1582, September 1962. (Cited on pages 23, 55 and 138.)
- [82] M S Steinberg. Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science*, 141(3579):401–408, August 1963. (Cited on pages 23 and 55.)
- [83] M S Steinberg. Does differential adhesion govern self-assembly processes in histogenesis. Equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *Journal of Experimental Zoology*, 173(4):395–433, April 1970. (Cited on pages 23 and 55.)
- [84] Thomas Lecuit and Pierre-François Lenne. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nature reviews Molecular cell biology*, 8(8):633–644, August 2007. (Cited on page 23.)
- [85] Jos Käfer, Takashi Hayashi, Athanasius F M Marée, Richard W Carthew, and François Graner. Cell adhesion and cortex contractility determine cell patterning in the Drosophila retina. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47):18549–18554, November 2007. (Cited on page 23.)
- [86] M Lisa Manning, Ramsey A Foty, Malcolm S Steinberg, and Eva-Maria Schoetz. Coaction of intercellular adhesion and cortical tension specifies tissue surface tension. *Proceedings of the National Academy of Sciences of the United States of America*, 107(28):12517–12522, July 2010. (Cited on pages 23 and 154.)
- [87] J L Maitre, H Berthoumieux, S F G Krens, G Salbreux, F Jülicher, E Paluch, and C-P Heisenberg. Adhesion Functions in Cell Sorting by Mechanically Coupling the Cortices of Adhering Cells. *Science*, 338(6104):253–256, October 2012. (Cited on pages 23 and 57.)
- [88] Gerard Apodaca, Luciana I Gallo, and David M Bryant. Role of membrane traffic in the generation of epithelial cell asymmetry. *Nature Cell Biology*, 14(12):1235–1243, December 2012. (Cited on page 24.)
- [89] John Philip Trinkaus. *Cells into organs*. the forces that shape the embryo. Prentice Hall, 1984. (Cited on page 24.)
- [90] Cornelis J Weijer. Collective cell migration in development. *Journal of Cell Science*, 122(Pt 18):3215–3223, September 2009. (Cited on page 24.)
- [91] Olga Ilina and Peter Friedl. Mechanisms of collective cell migration at a glance. *Journal of Cell Science*, 122(Pt 18):3203–3208, September 2009. (Cited on page 24.)
- [92] Pernille Rørth. Collective Cell Migration. *Annual review of cell and developmental biology*, 25(1):407–429, November 2009. (Cited on page 24.)

- [93] Sri Ram Krishna Vedula, Man Chun Leong, Tan Lei Lai, Pascal Hersen, Alexandre J Kabla, Chwee Teck Lim, and Benoît Ladoux. Emerging modes of collective cell migration induced by geometrical constraints. *Proceedings of the National Academy of Sciences*, 109(32):12974–12979, August 2012. (Cited on page 24.)
- [94] Michelangelo von Dassow and Lance A Davidson. Physics and the canalization of morphogenesis: a grand challenge in organismal biology. *Physical Biology*, 8(4):045002, July 2011. (Cited on page 24.)
- [95] J. D. Murray. *Mathematical biology II: spatial models and biomedical applications*. Springer, 2003. (Cited on pages 26, 50 and 73.)
- [96] Victor D Varner and Larry A Taber. On integrating experimental and theoretical models to determine physical mechanisms of morphogenesis. *Biosystems*, 109(3):412–419, September 2012. (Cited on page 27.)
- [97] Mathieu Ducros, Marcel van’t Hoff, Alexis Evrard, Christian Seebacher, Elke M Schmidt, Serge Charpak, and Martin Oheim. Efficient large core fiber-based detection for multi-channel two-photon fluorescence microscopy and spectral unmixing. *Journal of Neuroscience Methods*, 198(2):172–180, June 2011. (Cited on page 31.)
- [98] B Lombardot. *Description intégrative des cinématiques de la structure embryonnaire : morphogénèse précoce du poisson zébré*. PhD thesis, Ecole Polytechnique EDX Paris, October 2010. (Cited on pages 33, 175 and 176.)
- [99] Marvin Lee Minsky. Matter, mind and models. Technical report, Massachusetts Institute of Technology, 1965. (Cited on page 34.)
- [100] A M Turing. The chemical basis of morphogenesis. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, 237(641):37–72, 1952. (Cited on pages 37, 38, 50 and 64.)
- [101] N Metropolis. The beginning of the Monte Carlo method. *Los Alamos Science*, 15(584):125–130, 1987. (Cited on page 37.)
- [102] Katharina P Landsberg, Reza Farhadifar, Jonas Ranft, Daiki Umetsu, Thomas J Widmann, Thomas Bittig, Amani Said, Frank Jülicher, and Christian Dahmann. Increased Cell Bond Tension Governs Cell Sorting at the Drosophila Anteroposterior Compartment Boundary. *Current biology : CB*, 19(22):1950–1955, December 2009. (Cited on pages 39 and 57.)
- [103] M Behrndt, G Salbreux, P Campinho, R Hauschild, F Oswald, J Roensch, S W Grill, and C-P Heisenberg. Forces Driving Epithelial Spreading in Zebrafish Gastrulation. *Science*, 338(6104):257–260, October 2012. (Cited on pages 39, 152, 158, 161 and 195.)
- [104] Shin-Ichi Nishikawa, Lars Martin Jakt, and Takumi Era. Embryonic stem-cell culture as a tool for developmental cell biology. *Nature reviews Molecular cell biology*, 8(6):502–507, June 2007. (Cited on page 39.)
- [105] N N Oreskes, K K Shrader-Frechette, and K K Belitz. Verification, validation, and confirmation of numerical models in the Earth sciences. *Science*, 263(5147):641–646, February 1994. (Cited on pages 39 and 40.)
- [106] Popper. *The Logic of Scientific Discovery*. Basic Books, 1959. (Cited on page 39.)
- [107] Bas C van Fraassen. *The Scientific Image*. Oxford University Press, USA, reprint edition, October 1980. (Cited on page 40.)
- [108] Nicolas Rashevsky. *Mathematical Biophysics*. University of Chicago Press, 1938. (Cited on page 49.)
- [109] Paul Cull. The mathematical biophysics of Nicolas Rashevsky. *Biosystems*, 88(3):178–184, April 2007. (Cited on page 49.)

- [110] N Rashevsky. Topology and life: in search of general mathematical principles in biology and sociology. *Bulletin of Mathematical Biology*, 16(4):317–348, 1954. (Cited on page 49.)
- [111] N Rashevsky. An approach to the mathematical biophysics of biological self-regulation and of cell polarity. *Bulletin of Mathematical Biology*, 2(1):15–25, 1940. (Cited on page 49.)
- [112] A G Jacobson and R Gordon. Changes in the shape of the developing vertebrate nervous system analyzed experimentally, mathematically and by computer simulation. *Journal of Experimental Zoology*, 197(2):191–246, August 1976. (Cited on pages 49 and 50.)
- [113] A Gierer and H Meinhardt. A theory of biological pattern formation. *Biological Cybernetics*, 12(1):30–39, 1972. (Cited on pages 50 and 64.)
- [114] H Meinhardt and A Gierer. Applications of a theory of biological pattern formation based on lateral inhibition. *Journal of Cell Science*, 15(2):321–346, July 1974. (Cited on page 50.)
- [115] David A Young. A local activator-inhibitor model of vertebrate skin patterns. *Mathematical Biosciences*, 72(1):51–58, November 1984. (Cited on pages 50, 51 and 52.)
- [116] Keith Amonlirdviman, Narmada A Khare, David R P Tree, Wei-Shen Chen, Jeffrey D Axelrod, and Claire J Tomlin. Mathematical modeling of planar cell polarity to understand domineering nonautonomy. *Science*, 307(5708):423–426, January 2005. (Cited on page 50.)
- [117] Anne Karine Lagendijk, Andras Szabo, Roeland M H Merks, and Jeroen Bakkers. Hyaluronan A critical regulator of endothelial-to-mesenchymal transition during cardiac valve formation. *Trends in Cardiovascular Medicine*, pages 1–8, January 2013. (Cited on page 50.)
- [118] Shigeru Kondo and Takashi Miura. Reaction-diffusion model as a framework for understanding biological pattern formation. *Science*, 329(5999):1616–1620, September 2010. (Cited on page 50.)
- [119] J. D. Murray. *Mathematical Biology: I. An Introduction*. Springer, December 2002. (Cited on pages 50 and 73.)
- [120] Jonathan Bieler, Christian Pozzorini, and Felix Naef. Whole-embryo modeling of early segmentation in *Drosophila* identifies robust and fragile expression domains. *Biophysical Journal*, 101(2):287–296, July 2011. (Cited on page 51.)
- [121] Masafumi Inaba, Hiroaki Yamanaka, and Shigeru Kondo. Pigment pattern formation by contact-dependent depolarization. *Science*, 335(6069):677, February 2012. (Cited on page 51.)
- [122] Hans Meinhardt. *Models of biological pattern formation*. Academic Press, London, 1982. (Cited on pages 52 and 64.)
- [123] L Wolpert. Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, 25(1):1, 1969. (Cited on pages 52, 53 and 64.)
- [124] L Wolpert. Positional information revisited. *Development*, 107 Suppl:3–12, 1989. (Cited on page 52.)
- [125] L Wolpert. One hundred years of positional information. *Trends in genetics : TIG*, 12(9):359–364, September 1996. (Cited on page 52.)
- [126] Lewis Wolpert. Positional information and patterning revisited. *Journal of Theoretical Biology*, 269(1):359–365, January 2011. (Cited on page 52.)
- [127] P A Lawrence. Morphogens: how big is the big picture? *Nature Cell Biology*, 3(7):E151–4, July 2001. (Cited on page 53.)
- [128] Arthur D Lander, Qing Nie, and Frederic Y M Wan. Do morphogen gradients arise by diffusion? *Developmental cell*, 2(6):785–796, June 2002. (Cited on page 53.)
- [129] Tetsuya Tabata and Yuki Takei. Morphogens, their identification and regulation. *Development*, 131(4):703–712, February 2004. (Cited on page 53.)

- [130] Anna Kicheva, Tobias Bollenbach, Ortrud Wartlick, Frank Jülicher, and Marcos González-Gaitán. Investigating the principles of morphogen gradient formation: from tissues to cells. *Current opinion in genetics & development*, pages 1–6, September 2012. (Cited on page 53.)
- [131] P Eggenberger. Evolving morphologies of simulated 3D organisms based on differential gene expression. *Proceedings of the Fourth European Conference on Artificial Life*, pages 205–213, 1997. (Cited on page 53.)
- [132] S Kumar and P Bentley. Biologically inspired evolutionary development. *Evolvable Systems: From Biology to Hardware*, pages 57–68, 2003. (Cited on page 53.)
- [133] R Doursat. The Growing Canvas of Biological Development: Multiscale Pattern Generation on an Expanding Lattice of Gene Regulatory Nets. *InterJournal: Complex Systems 1809*, February 2006. (Cited on page 53.)
- [134] JF Miller and W Banzhaf. Evolving the program for a cell: from french flags to boolean circuits. *On Growth, Form and Computers*, pages 278–301, 2003. (Cited on page 53.)
- [135] J Knabe, M Schilstra, and C L Nehaniv. Evolution and morphogenesis of differentiated multicellular organisms: autonomously generated diffusion gradients for positional information. *Artificial Life XI*, 2008. (Cited on page 53.)
- [136] N Barkai and B-Z Shilo. Robust Generation and Decoding of Morphogen Gradients. *Cold Spring Harbor Perspectives in Biology*, 1(5):a001990–a001990, November 2009. (Cited on page 53.)
- [137] S Kauffman. Homeostasis and differentiation in random genetic control networks. *Nature*, 224(5215):177–178, October 1969. (Cited on page 54.)
- [138] J Reinitz and DH Sharp. Mechanism of eve stripe formation. *Mechanisms of Development*, 49(1):133, 1995. (Cited on page 54.)
- [139] A Rolland-Lagan, M Paquette, V Tweedle, and M Akimenko. Morphogen-based simulation model of ray growth and joint patterning during fin development and regeneration. *Development*, 139:1–10, February 2012. (Cited on page 54.)
- [140] Danny Ben-Zvi, Ben-Zion Shilo, and Naama Barkai. Scaling of morphogen gradients. *Current opinion in genetics & development*, 21(6):704–710, December 2011. (Cited on page 54.)
- [141] Peter McHale, Wouter-Jan Rappel, and Herbert Levine. Embryonic pattern scaling achieved by oppositely directed morphogen gradients. *Physical Biology*, 3(2):107–120, June 2006. (Cited on page 54.)
- [142] Thomas Gregor, William Bialek, Rob R de Ruyter van Steveninck, David W Tank, and Eric F Wieschaus. Diffusion and scaling during early embryonic pattern formation. *Proceedings of the National Academy of Sciences of the United States of America*, 102(51):18403–18407, December 2005. (Cited on page 54.)
- [143] H G Othmer and E Pate. Scale-invariance in reaction-diffusion models of spatial pattern formation. *Proceedings of the National Academy of Sciences of the United States of America*, 77(7):4180–4184, July 1980. (Cited on page 54.)
- [144] Matthew C Gibson, Ankit B Patel, Radhika Nagpal, and Norbert Perrimon. The emergence of geometric order in proliferating metazoan epithelia. *Nature*, 442(7106):1038–1041, August 2006. (Cited on pages 54, 55, 56, 251 and 253.)
- [145] R Farhadifar, JC Röper, B Aigouy, S Eaton, and F Jülicher. The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. *Current Biology*, 17(24):2095–2104, 2007. (Cited on pages 54, 56, 64 and 80.)
- [146] Ankit B Patel, William T Gibson, Matthew C Gibson, and Radhika Nagpal. Modeling and inferring cleavage patterns in proliferating epithelia. *PLoS Computational Biology*, 5(6):e1000412–e1000412, June 2009. (Cited on page 55.)

- [147] Patrik Sahlin and Henrik Jönsson. A modeling study on how cell division affects properties of epithelial tissues under isotropic growth. *PLoS ONE*, 5(7):e11750, 2010. (Cited on pages 55, 64 and 80.)
- [148] SA Sandersius, M Chuai, CJ Weijer, and TJ Newman. Correlating Cell Behavior with Tissue Topology in Embryonic Epithelia. *PLoS ONE*, 6(4):e18081, 2011. (Cited on pages 55, 64, 82, 251 and 253.)
- [149] Luis M Escudero, Luciano da F Costa, Anna Kicheva, James Briscoe, Matthew Freeman, and M Madan Babu. Epithelial organisation revealed by a network of cellular contacts. *Nature communications*, 2:526, 2011. (Cited on pages 55 and 248.)
- [150] William T Gibson, James H Veldhuis, Boris Rubinstein, Heather N Cartwright, Norbert Perrimon, G Wayne Brodland, Radhika Nagpal, and Matthew C Gibson. Control of the mitotic cleavage plane by local epithelial topology. *Cell*, 144(3):427–438, February 2011. (Cited on page 55.)
- [151] M S Steinberg. Mechanism of tissue reconstruction by dissociated cells. II. Time-course of events. *Science*, 137(3532):762–763, September 1962. (Cited on page 55.)
- [152] M S Steinberg. On the mechanism of tissue reconstruction by dissociated cells, III. Free energy relations and the reorganization of fused, heteronomic tissue fragments. *Proceedings of the National Academy of Sciences of the United States of America*, 48(10):1769–1776, October 1962. (Cited on page 55.)
- [153] F Graner and JA Glazier. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Physical Review Letters*, 69(13):2013–2016, September 1992. (Cited on pages 57, 77 and 138.)
- [154] James Glazier and François Graner. Simulation of the differential adhesion driven rearrangement of biological cells. *Physical Review E*, 47(3):2128–2154, March 1993. (Cited on pages 57, 64 and 77.)
- [155] G W Brodland and H H Chen. The mechanics of heterotypic cell aggregates: insights from computer simulations. *Journal of biomechanical engineering*, 122(4):402–407, August 2000. (Cited on page 57.)
- [156] Maryam Aliee, Jens-Christian Röper, Katharina P Landsberg, Constanze Pentzold, Thomas J Widmann, Frank Jülicher, and Christian Dahmann. Physical Mechanisms Shaping the Drosophila Dorsoventral Compartment Boundary. *Current biology : CB*, pages 1–10, May 2012. (Cited on page 57.)
- [157] Carine P Beatrice and Leonardo G Brunnet. Cell sorting based on motility differences. *Physical Review E*, 84(3 Pt 1):031927, September 2011. (Cited on page 57.)
- [158] RA Foty and MS Steinberg. The differential adhesion hypothesis: a direct evaluation. *Developmental Biology*, 278(1):255–263, 2005. (Cited on page 57.)
- [159] Y Zhang, G L Thomas, M Swat, A Shirinifard, and J A Glazier. Computer Simulations of Cell Sorting Due to Differential Adhesion. *PLoS ONE*, 6(10):e24999, 2011. (Cited on pages 57, 77, 128 and 138.)
- [160] Sean B Carroll. *Endless Forms Most Beautiful*. The New Science of Evo Devo. W. W. Norton & Company, April 2006. (Cited on page 57.)
- [161] Dr Marc W Kirschner, John C Gerhart, and John Norton. *The Plausibility of Life*. Resolving Darwin’s Dilemma. Yale University Press, 2005. (Cited on page 58.)
- [162] Stuart A Kauffman. *The Origins of Order*. Self Organization and Selection in Evolution. Oxford University Press, USA, 1993. (Cited on page 58.)
- [163] Brian C Goodwin. *How the Leopard Changed Its Spots*. The Evolution of Complexity. Princeton University Press, 1994. (Cited on page 58.)

- [164] Guy B Blanchard and Richard J Adams. Measuring the multi-scale integration of mechanical forces during morphogenesis. *Current opinion in genetics & development*, pages 1–11, September 2011. (Cited on pages 58 and 152.)
- [165] I M M van Leeuwen, G R Mirams, A Walter, A Fletcher, P Murray, J Osborne, S Varma, S J Young, J Cooper, B Doyle, J Pitt-Francis, L Momtahan, P Pathmanathan, J P Whiteley, S J Chapman, D J Gavaghan, O E Jensen, J R King, P K Maini, S L Waters, and H M Byrne. An integrative computational model for intestinal tissue renewal. *Cell Proliferation*, 42(5):617–636, October 2009. (Cited on page 58.)
- [166] Ramon Grima. Multiscale modeling of biological pattern formation. *Current topics in developmental biology*, 81:435–460, 2008. (Cited on page 58.)
- [167] Shawn C Little and Eric F Wieschaus. Shifting patterns: merging molecules, morphogens, motility, and methodology. *Developmental cell*, 21(1):2–4, July 2011. (Cited on page 58.)
- [168] JA Izaguirre, R Chaturvedi, C Huang, T Cickovski, J Coffland, G Thomas, G Forgacs, M Alber, G Hentschel, and SA Newman. CompuCell, a multi-model framework for simulation of morphogenesis. *Bioinformatics*, 20(7):1129–1137, 2004. (Cited on pages 58 and 77.)
- [169] T Cickovski, K Aras, M Swat, RMH Merks, T Glimm, HGE Hentschel, MS Alber, JA Glazier, SA Newman, and JA Izaguirre. From genes to organisms via the cell: a problem-solving environment for multicellular development. *Computing in Science & Engineering*, 9(4):50–60, 2007. (Cited on pages 58 and 77.)
- [170] Maciej H MH Swat, Susan D SD Hester, Ariel I AI Balter, Randy W RW Heiland, Benjamin L BL Zaitlen, and James A JA Glazier. Multicell simulations of development and disease using the CompuCell3D simulation environment. *Methods in molecular biology (Clifton, NJ)*, 500:361–428, December 2008. (Cited on pages 58 and 77.)
- [171] R Doursat. Programmable architectures that are complex and self-organized: From morphogenesis to engineering. *Proceedings of Artificial XI*, pages 181–188, 2008. (Cited on page 58.)
- [172] M Joachimczak and B Wróbel. Evo-devo in silico: a model of a gene network regulating multicellular development in 3D space with artificial physics. *Artificial Life XI: Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*, pages 297–304, 2008. (Cited on page 58.)
- [173] L Schramm, Y Jin, and B Sendhoff. Emerged coupling of motor control and morphological development in evolution of multi-cellular animats. *Advances in Artificial Life. Darwin Meets von Neumann*, pages 27–34, 2011. (Cited on page 58.)
- [174] René Doursat, Carlos Sánchez, Razvan Dordea, David Fourquet, and Taras Kowaliw. Embryomorphic Engineering: Emergent innovation through evolutionary development. In René Doursat, Hiroki Sayama, and Olivier Michel, editors, *Morphogenetic Engineering: Toward Programmable Complex Systems*. Springer, February 2012. (Cited on page 58.)
- [175] R P Feynman, R B Leighton, and M L Sands. The Feynman Lectures on Physics, vol I. Addison-Wesley, 1964. (Cited on page 60.)
- [176] P W Anderson. More is different. *Science*, 1972. (Cited on page 60.)
- [177] M O Jensen, V Jogini, D W Borhani, A E Leffler, R O Dror, and D E Shaw. Mechanism of Voltage Gating in Potassium Channels. *Science*, 336(6078):229–233, April 2012. (Cited on page 61.)
- [178] S Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995. (Cited on page 61.)
- [179] Philip J Murray, Philip K Maini, and Ruth E Baker. The clock and wavefront model revisited. *Journal of Theoretical Biology*, 283(1):227–238, August 2011. (Cited on page 62.)

- [180] A R A Anderson, M A J Chaplain, and K A Rejniak. *Single-cell-based Models in Biology and Medicine*. Birkhäuser Basel; Boston, 2007. (Cited on page 63.)
- [181] M Weliky and G Oster. The mechanical basis of cell rearrangement. I. Epithelial morphogenesis during Fundulus epiboly. *Development*, 109(2):373–386, June 1990. (Cited on pages 64 and 80.)
- [182] M Weliky, S Minsuk, R Keller, and G Oster. Notochord morphogenesis in Xenopus laevis: simulation of cell behavior underlying tissue convergence and extension. *Development*, 113(4):1231–1244, December 1991. (Cited on pages 64, 80 and 298.)
- [183] T J Newman. Modeling multicellular systems using subcellular elements. *Mathematical biosciences and engineering : MBE*, 2(3):613–624, July 2005. (Cited on pages 64, 81 and 297.)
- [184] S A Sandersius, M Chuai, C J Weijer, and T J Newman. A ‘chemotactic dipole’ mechanism for large-scale vortex motion during primitive streak formation in the chick embryo. *Physical Biology*, 8(4):045008, August 2011. (Cited on pages 64, 81, 82, 295 and 298.)
- [185] G Schaller and M Meyer-Hermann. Kinetic and dynamic Delaunay tetrahedralizations in three dimensions. *Computer Physics Communications*, 162(1):9–23, 2004. (Cited on pages 64 and 79.)
- [186] Vincent Fleury. An elasto-plastic model of avian gastrulation. *Organogenesis*, 2(1):6–16, January 2005. (Cited on pages 64, 73 and 74.)
- [187] V Fleury. Dynamic topology of the cephalochordate to amniote morphological transition: A self-organized system of Russian dolls. *Comptes Rendus Biologies*, 2011. (Cited on page 64.)
- [188] William J R Longabaugh, Eric H Davidson, and Hamid Bolouri. Computational representation of developmental genetic regulatory networks. *Developmental Biology*, 283(1):1–16, July 2005. (Cited on pages 64 and 114.)
- [189] Sriram Ramaswamy. The Mechanics and Statistics of Active Matter. *Annual Review of Condensed Matter Physics*, 1(1):323–345, August 2010. (Cited on page 71.)
- [190] J. D. Murray and G F Oster. Generation of biological pattern and form. *IMA journal of mathematics applied in medicine and biology*, 1(1):51–75, 1984. (Cited on pages 73 and 75.)
- [191] Olivier Hamant, Marcus G Heisler, Henrik Jönsson, Pawel Krupinski, Magalie Uyttewaal, Plamen Bokov, Francis Corson, Patrik Sahlin, Arezki Boudaoud, Elliot M Meyerowitz, Yves Couder, and Jan Traas. Developmental patterning by mechanical signals in Arabidopsis. *Science*, 322(5908):1650–1655, December 2008. (Cited on pages 74 and 81.)
- [192] Vito Conte, José J Muñoz, and Mark Miodownik. A 3D finite element model of ventral furrow invagination in the *Drosophila melanogaster* embryo. *Journal of the mechanical behavior of biomedical materials*, 1(2):188–198, April 2008. (Cited on page 74.)
- [193] G Forgacs, RA Foty, Y Shafir, and MS Steinberg. Viscoelastic properties of living embryonic tissues: a quantitative study. *Biophysical Journal*, 74(5):2227–2234, 1998. (Cited on page 75.)
- [194] S R Lubkin and Z Li. Force and deformation on branching rudiments: cleaving between hypotheses. *Biomechanics and Modeling in Mechanobiology*, 1(1):5–16, June 2002. (Cited on page 76.)
- [195] Philippe-Alexandre Pouille and Emmanuel Farge. Hydrodynamic simulation of multicellular embryo invagination. *Physical Biology*, 5(1):15005, 2008. (Cited on page 76.)
- [196] V Fleury. A change in boundary conditions induces a discontinuity of tissue flow in chicken embryos and the formation of the cephalic fold. *The European physical journal E, Soft matter*, 34(7):73, July 2011. (Cited on page 76.)
- [197] Andreas Deutsch and Sabine Dormann. *Cellular Automaton Modeling of Biological Pattern Formation*. Characterization, Applications, and Analysis. Birkhauser Boston, September 2005. (Cited on page 76.)

- [198] A Anderson. A hybrid multiscale model of solid tumour growth and invasion: Evolution and the microenvironment. *Single-Cell-Based Models in Biology and Medicine*, pages 3–28, 2007. (Cited on page 76.)
- [199] Haralampos Hatzikirou and Andreas Deutsch. Cellular automata as microscopic models of cell migration in heterogeneous environments. *Current topics in developmental biology*, 81:401–434, 2008. (Cited on page 76.)
- [200] David Peak, Jevin D West, Susanna M Messinger, and Keith A Mott. Evidence for complex, collective dynamics and emergent, distributed computation in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 101(4):918–922, January 2004. (Cited on page 76.)
- [201] J A Glazier, A Balter, and N J Popławski. Magnetization to morphogenesis: a brief history of the Glazier-Graner-Hogeweg model. *Single-Cell-Based Models in Biology and Medicine*, pages 79–106, 2007. (Cited on page 77.)
- [202] A Balter, R M H Merks, N J Popławski, M Swat, and J A Glazier. The Glazier-Graner-Hogeweg model: extensions, future directions, and opportunities for further study. *Single-Cell-Based Models in Biology and Medicine*, pages 151–167, 2007. (Cited on page 77.)
- [203] A F M Marée, V A Grieneisen, and P Hogeweg. The Cellular Potts Model and biophysical properties of cells, tissues and morphogenesis. *Single-Cell-Based Models in Biology and Medicine*, pages 107–136, 2007. (Cited on page 77.)
- [204] N J Savill and R M H Merks. The cellular Potts model in biomedicine. *Single-Cell-Based Models in Biology and Medicine*, pages 137–150, 2007. (Cited on page 77.)
- [205] AFM Marée and P Hogeweg. How amoeboids self-organize into a fruiting body: multicellular coordination in *Dictyostelium discoideum*. *Proceedings of the National Academy of Sciences*, 98(7):3879, March 2001. (Cited on page 77.)
- [206] Nigel C Harrison, Ruth Diez Del Corral, and Bakhtier Vasiev. Coordination of cell differentiation and migration in mathematical models of caudal embryonic axis extension. *PLoS ONE*, 6(7):e22700, 2011. (Cited on page 77.)
- [207] James A Glazier, Ying Zhang, Maciej Swat, Benjamin Zaitlen, and Santiago Schnell. Coordinated action of N-CAM, N-cadherin, EphA4, and ephrinB2 translates genetic prepatterns into structure during somitogenesis in chick. *Current topics in developmental biology*, 81:205–247, 2008. (Cited on page 77.)
- [208] Susan D Hester, Julio M Belmonte, J Scott Gens, Sherry G Clendenon, and James A Glazier. A Multi-cell, Multi-scale Model of Vertebrate Segmentation and Somite Formation. *PLoS Computational Biology*, 7(10):e1002155, October 2011. (Cited on page 77.)
- [209] Roeland M H Merks, Sergey V Brodsky, Michael S Goligorsky, Stuart A Newman, and James A Glazier. Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling. *Developmental Biology*, 289(1):44–54, January 2006. (Cited on page 77.)
- [210] D Drasdo, R Kree, and JS McCaskill. Monte Carlo approach to tissue-cell populations. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 52(6):6635–6657, December 1995. (Cited on page 79.)
- [211] E Palsson. A three-dimensional model of cell movement in multicellular systems. *Future Generation Computer Systems*, 17(7):835–852, 2001. (Cited on pages 79 and 305.)
- [212] T Beyer, G Schaller, A Deutsch, and M Meyer-Hermann. Parallel dynamic and kinetic regular triangulation in three dimensions. *Computer Physics Communications*, 172(2):86–108, 2005. (Cited on page 79.)

- [213] Michael Meyer-Hermann. Delaunay-Object-Dynamics: cell mechanics with a 3D kinetic and dynamic weighted Delaunay-triangulation. *Current topics in developmental biology*, 81:373–399, 2008. (Cited on pages 79, 297 and 298.)
- [214] G Schaller and M Meyer-Hermann. Epidermal homeostasis control in an off-lattice agent-based model. *Arxiv preprint physics/0507059*, 2005. (Cited on page 79.)
- [215] T Beyer and M Meyer-Hermann. Modeling emergent tissue organization involving high-speed migrating cells in a flow equilibrium. *Physical Review E*, 76(2):021929, 2007. (Cited on page 79.)
- [216] Masataka Yamao, Honda Naoki, and Shin Ishii. Multi-Cellular Logistics of Collective Cell Migration. *PLoS ONE*, 6(12):e27950, December 2011. (Cited on pages 79, 106 and 298.)
- [217] D E Ingber. Tensegrity I. Cell structure and hierarchical systems biology. *Journal of Cell Science*, 116(7):1157–1173, April 2003. (Cited on page 79.)
- [218] D E Ingber. Tensegrity II. How structural networks influence cellular information processing networks. *Journal of Cell Science*, 116(8):1397–1408, April 2003. (Cited on page 79.)
- [219] G M Odell, G Oster, P Alberch, and B Burnside. The mechanical basis of morphogenesis. I. Epithelial folding and invagination. *Developmental Biology*, 85(2):446–462, July 1981. (Cited on pages 80, 91, 132 and 297.)
- [220] Hisao Honda, Nami Motosugi, Tatsuzo Nagai, Masaharu Tanemura, and Takashi Hiiragi. Computer simulation of emerging asymmetry in the mouse blastocyst. *Development*, 135(8):1407–1414, April 2008. (Cited on page 81.)
- [221] Kristin Sherrard, François Robin, Patrick Lemaire, and Edwin Munro. Sequential Activation of Apical and Basolateral Contractility Drives Ascidian Endoderm Invagination. *Current Biology*, 20(17):1499–1510, September 2010. (Cited on pages 81 and 297.)
- [222] Oriol Canela-Xandri, Francesc Sagués, Jaume Casademunt, and Javier Buceta. Dynamics and Mechanical Stability of the Developing Dorsoventral Organizer of the Wing Imaginal Disc. *PLoS Computational Biology*, 7(9):e1002153, September 2011. (Cited on page 81.)
- [223] S A Sandersius, C J Weijer, and T J Newman. Emergent cell and tissue dynamics from subcellular modeling of active biomechanical processes. *Physical Biology*, 8(4):045007, August 2011. (Cited on pages 81 and 82.)
- [224] Sebastian A Sandersius and Timothy J Newman. Modeling cell rheology with the Subcellular Element Model. *Physical Biology*, 5(1):15002, 2008. (Cited on pages 81, 82 and 295.)
- [225] John L Semple, Nicholis Woolridge, and Charles J Lumsden. In vitro, in vivo, in silico: computational systems in tissue engineering and regenerative medicine. *Tissue engineering*, 11(3-4):341–356, March 2005. (Cited on page 82.)
- [226] D Gonzalez-Rodriguez, K Guevorkian, S Douezan, and F Brochard-Wyart. Soft Matter Models of Developing Tissues and Tumors. *Science*, 338(6109):910–917, November 2012. (Cited on page 82.)
- [227] Matthew A Wyczalkowski, Zi Chen, Benjamen A Filas, Victor D Varner, and Larry A Taber. Computational models for mechanics of morphogenesis. *Birth Defects Research Part C: Embryo Today: Reviews*, 96(2):132–152, June 2012. (Cited on page 82.)
- [228] Andreea Robu, Roxana Aldea, Oana Munteanu, Monica Neagu, Lacramioara Stoicu-Tivadar, and Adrian Neagu. Computer simulations of in vitro morphogenesis. *Biosystems*, 109(3):430–443, September 2012. (Cited on page 82.)
- [229] KL Johnson, K Kendall, and AD Roberts. Surface energy and the contact of elastic solids. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 324(1558):301, 1971. (Cited on page 86.)

- [230] B V Derjaguin, V M Muller, and Yu P Toporov. Effect of contact deformations on the adhesion of particles. *Journal of Colloid and Interface Science*, 53(2):314–326, November 1975. (Cited on page 87.)
- [231] RM Purcell. Life at low Reynolds number. *American Journal of Physics*, 1977. (Cited on pages 90 and 297.)
- [232] C Claycomb and J Tran. *Introductory Biophysics: Perspectives on the Living State*. John and Bartlett Publishers, LLC, 2010. (Cited on page 90.)
- [233] Dominique Attali and Annick Montanvert. Computing and simplifying 2D and 3D continuous skeletons. *Computer Vision and Image Understanding*, 1997. (Cited on page 98.)
- [234] F Cazals and J Giesen. Delaunay Triangulation Based Surface Reconstruction : Ideas and Algorithms. *INRIA Rapport de recherche*, pages 1–45, November 2004. (Cited on page 99.)
- [235] Bartek Wilczynski and Eileen E M Furlong. Challenges for modeling global gene regulatory networks during development: Insights from Drosophila. *Developmental Biology*, 340(2):161–169, April 2010. (Cited on page 113.)
- [236] S Ben-Tabou de Leon and EH Davidson. Modeling the dynamics of transcriptional gene regulatory networks for animal development. *Developmental Biology*, 325(2):317–328, 2009. (Cited on page 113.)
- [237] Clare E Giacomantonio and Geoffrey J Goodhill. A Boolean Model of the Gene Regulatory Network Underlying Mammalian Cortical Area Development. *PLoS Computational Biology*, 6(9):e1000936, September 2010. (Cited on page 113.)
- [238] William J R Longabaugh, Eric H Davidson, and Hamid Bolouri. Visualization, documentation, analysis, and communication of large-scale gene regulatory networks. *BBA - Gene Regulatory Mechanisms*, 1789(4):363–374, April 2009. (Cited on page 114.)
- [239] Isabelle S Peter, Emmanuel Faure, and Eric H Davidson. Predictive computation of genomic logic processing functions in embryonic development. *Proceedings of the National Academy of Sciences*, 109(41):16434–16442, October 2012. (Cited on pages 118 and 300.)
- [240] Nikola Kojić, Austin Huang, Euiheon Chung, Miloš Ivanović, Nenad Filipović, Miloš Kojić, and Daniel J Tschumperlin. A 3-D Model of Ligand Transport in a Deforming Extracellular Space. *Biophysj*, 99(11):3517–3525, December 2010. (Cited on page 120.)
- [241] Eric H Davidson. Emerging properties of animal gene regulatory networks. *Nature*, 468(7326):911–920, December 2010. (Cited on page 121.)
- [242] Isabelle S Peter and Eric H Davidson. Modularity and design principles in the sea urchin embryo gene regulatory network. *FEBS letters*, 583(24):3948–3958, December 2009. (Cited on page 121.)
- [243] Eric H Davidson. Network design principles from the sea urchin embryo. *Current opinion in genetics & development*, 19(6):535–540, December 2009. (Cited on page 121.)
- [244] R R Thomas. Boolean formalization of genetic control circuits. *Journal of Theoretical Biology*, 42(3):563–585, November 1973. (Cited on page 126.)
- [245] Titus J Boggon, John Murray, Sophie Chappuis-Flament, Ellen Wong, Barry M Gumbiner, and Lawrence Shapiro. C cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science*, 296(5571):1308–1313, May 2002. (Cited on page 128.)
- [246] Julien G Dumortier, Sandro Martin, Dirk Meyer, Frédéric M Rosa, and Nicolas B David. Collective mesendoderm migration relies on an intrinsic directionality signal transmitted through cell contacts. *Proceedings of the National Academy of Sciences*, 109(42):16945–16950, October 2012. (Cited on page 131.)

- [247] L A D'Amico and M S Cooper. Morphogenetic domains in the yolk syncytial layer of axiating zebrafish embryos. *Developmental dynamics : an official publication of the American Association of Anatomists*, 222(4):611–624, December 2001. (Cited on pages 151, 162, 163 and 165.)
- [248] Juan-Antonio Montero and Carl-Philipp Heisenberg. Gastrulation dynamics: cells move into focus. *Trends in cell biology*, 14(11):620–627, November 2004. (Cited on pages 151 and 162.)
- [249] Lilianna Solnica-Krezel. Conserved Patterns of Cell Movements during Vertebrate Gastrulation. *Current Biology*, 15(6):R213–R228, March 2005. (Cited on page 151.)
- [250] Lilianna Solnica-Krezel. Gastrulation in zebrafish – all just about adhesion? *Current opinion in genetics & development*, 16(4):433–441, August 2006. (Cited on page 151.)
- [251] Laurel A Rohde and Carl-Philipp Heisenberg. Zebrafish gastrulation: cell movements, signals, and mechanisms. *International review of cytology*, 261:159–192, 2007. (Cited on pages 151 and 155.)
- [252] Carl-Philipp Heisenberg and Lilianna Solnica-Krezel. Back and forth between cell fate specification and movement during vertebrate gastrulation. *Current opinion in genetics & development*, 18(4):311–316, August 2008. (Cited on page 151.)
- [253] C B Kimmel, W W Ballard, S R Kimmel, B Ullmann, and T F Schilling. Stages of embryonic development of the zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists*, 203(3):253–310, July 1995. (Cited on pages 153, 169, 174, 228 and 263.)
- [254] Nicolas Olivier, Miguel A Luengo-Oroz, Louise Duloquin, Emmanuel Faure, Thierry Savy, Israël Veilleux, Xavier Solinas, Delphine Débarre, Paul Bourgine, Andrés Santos, Nadine Peyriéras, and Emmanuel Beaurepaire. Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. *Science*, 329(5994):967–971, August 2010. (Cited on pages 152, 196, 211, 212, 214, 228 and 240.)
- [255] D A Kane, R M Warga, and C B Kimmel. Mitotic domains in the early embryo of the zebrafish. *Nature*, 360(6406):735–737, December 1992. (Cited on page 154.)
- [256] C B Kimmel and R D Law. Cell lineage of zebrafish blastomeres. III. Clonal analyses of the blastula and gastrula stages. *Developmental Biology*, 108(1):94–101, March 1985. (Cited on pages 154 and 155.)
- [257] CB Kimmel. Origin and organization of the zebrafish fate map. *Development*, 108(4):581–594, 1990. (Cited on pages 154 and 229.)
- [258] S F Gabby Krens, Stephanie Möllmert, and Carl-Philipp Heisenberg. Enveloping cell-layer differentiation at the surface of zebrafish germ-layer tissue explants. *Proceedings of the National Academy of Sciences*, 108(3):E9–10– author reply E11, January 2011. (Cited on page 154.)
- [259] Sung-Kook Hong, Carly S Levin, Jamie L Brown, Haiyan Wan, Brad T Sherman, Da Wei Huang, Richard A Lempicki, and Benjamin Feldman. Pre-gastrula expression of zebrafish extraembryonic genes. *BMC developmental biology*, 10:42, 2010. (Cited on page 154.)
- [260] C B Kimmel and R D Law. Cell lineage of zebrafish blastomeres. II. Formation of the yolk syncytial layer. *Developmental Biology*, 108(1):86–93, March 1985. (Cited on page 155.)
- [261] Lilianna Solnica-Krezel. *Pattern Formation in Zebrafish*. Springer, 2002. (Cited on page 155.)
- [262] L Solnica-Krezel and W Driever. Microtubule arrays of the zebrafish yolk cell: organization and function during epiboly. *Development*, 120(9):2443–2455, September 1994. (Cited on pages 155, 156, 158, 159, 160, 165 and 264.)

- [263] Michael Y F Yuen, Sarah E Webb, Ching Man Chan, Bernard Thisse, Christine Thisse, and Andrew L Miller. Characterization of Ca(2+) signaling in the external yolk syncytial layer during the late blastula and early gastrula periods of zebrafish development. *Biochimica et biophysica acta*, November 2012. (Cited on page 157.)
- [264] RD Fink and MS Cooper. Apical membrane turnover is accelerated near cell cell contacts in an embryonic epithelium. *Developmental Biology*, 174(2):180–189, 1996. (Cited on pages 155 and 245.)
- [265] R M Warga and C B Kimmel. Cell movements during epiboly and gastrulation in zebrafish. *Development*, 108(4):569–580, April 1990. (Cited on pages 155, 164 and 244.)
- [266] Trinkaus. A study of the mechanism of epiboly in the egg of Fundulus heteroclitus. *Journal of Experimental Zoology*, pages 1–51, 1951. (Cited on page 159.)
- [267] J P Trinkaus and T L Lentz. Surface specializations of Fundulus cells and their relation to cell movements during gastrulation. *The Journal of cell biology*, 32(1):139–153, January 1967. (Cited on page 159.)
- [268] J P Trinkaus. Mechanism of Fundulus epiboly - a current view. *Integrative and Comparative Biology*, pages 1–16, 1984. (Cited on pages 159 and 161.)
- [269] T Betchaku and J P Trinkaus. Contact relations, surface activity, and cortical microfilaments of marginal cells of the enveloping layer and of the yolk syncytial and yolk cytoplasmic layers of fundulus before and during epiboly. *Journal of Experimental Zoology*, 206(3):381–426, November 1978. (Cited on pages 159 and 160.)
- [270] T Betchaku and J P Trinkaus. Programmed endocytosis during epiboly of Fundulus heteroclitus. *American zoologist*, 26(1):193–199, 1986. (Cited on page 160.)
- [271] S E Zalik, E Lewandowski, Z Kam, and B Geiger. Cell adhesion and the actin cytoskeleton of the enveloping layer in the zebrafish embryo during epiboly. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, 77(6):527–542, 1999. (Cited on page 160.)
- [272] R E Keller and J P Trinkaus. Rearrangement of enveloping layer cells without disruption of the epithelial permeability barrier as a factor in Fundulus epiboly. *Developmental Biology*, 120(1):12–24, March 1987. (Cited on page 160.)
- [273] J Todd Blankenship, Stephanie T Backovic, Justina S P Sanny, Ori Weitz, and Jennifer A Zallen. Multicellular Rosette Formation Links Planar Cell Polarity to Tissue Morphogenesis. *Developmental cell*, 11(4):459–470, October 2006. (Cited on page 160.)
- [274] Mathias Köppen, Beatriz García Fernández, Lara Carvalho, Antonio Jacinto, and Carl-Philipp Heisenberg. Coordinated cell-shape changes control epithelial movement in zebrafish and Drosophila. *Development*, 133(14):2671–2681, July 2006. (Cited on pages 160 and 161.)
- [275] Jackie C Cheng, Andrew L Miller, and Sarah E Webb. Organization and function of microfilaments during late epiboly in zebrafish embryos. *Developmental dynamics : an official publication of the American Association of Anatomists*, 231(2):313–323, October 2004. (Cited on page 160.)
- [276] J P Trinkaus and C A Erickson. Protrusive activity, mode and rate of locomotion, and pattern of adhesion of Fundulus deep cells during gastrulation. *Journal of Experimental Zoology*, 228(1):41–70, October 1983. (Cited on page 161.)
- [277] Philipp J Keller, Annette D Schmidt, Joachim Wittbrodt, and Ernst H K Stelzer. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science*, 322(5904):1065–1069, November 2008. (Cited on pages 161, 164, 167, 177, 179, 286, 289 and 295.)
- [278] Juan-Antonio Montero, Lara Carvalho, Michaela Wilsch-Bräuninger, Beate Kilian, Chigdem Mustafa, and Carl-Philipp Heisenberg. Shield formation at the onset of zebrafish gastrulation. *Development*, 132(6):1187–1198, February 2005. (Cited on pages 161, 162 and 163.)

- [279] D A Kane. Mutations in half baked/E-cadherin block cell behaviors that are necessary for teleost epiboly. *Development*, 132(5):1105–1116, February 2005. (Cited on page 161.)
- [280] DA Kane. The zebrafish early arrest mutants. *Development*, 123(1):57–66, 1996. (Cited on pages 161 and 211.)
- [281] Lara Carvalho, Jan Stühmer, Justin S Bois, Yannis Kalaidzidis, Virginie Lecaudey, and Carl-Philipp Heisenberg. Control of convergent yolk syncytial layer nuclear movement in zebrafish. *Development*, 136(8):1305–1315, April 2009. (Cited on page 162.)
- [282] Florian Ulrich, Michael Krieg, Eva-Maria Schoetz, Vinzenz Link, Irinka Castanon, Viktor Schnabel, Anna Taubenberger, Daniel Mueller, Pierre-Henri Puech, and Carl-Philipp Heisenberg. Wnt11 functions in gastrulation by controlling cell cohesion through Rab5c and E-cadherin. *Developmental cell*, 9(4):555–564, October 2005. (Cited on page 162.)
- [283] J P Trinkaus. Ingression during early gastrulation of fundulus. *Developmental Biology*, 177(1):356–370, July 1996. (Cited on page 162.)
- [284] A Carmany-Rampey and A F Schier. Single-cell internalization during zebrafish gastrulation. *Current biology : CB*, 11(16):1261–1265, August 2001. (Cited on pages 163 and 292.)
- [285] N B David and F M Rosa. Cell autonomous commitment to an endodermal fate and behaviour by activation of Nodal signalling. *Development*, 128(20):3937–3947, October 2001. (Cited on page 164.)
- [286] T Mizoguchi, H Verkade, J K Heath, A Kuroiwa, and Y Kikuchi. Sdf1 Cxcr4 signaling controls the dorsal migration of endodermal cells during zebrafish gastrulation. *Development*, 135(15):2521–2529, June 2008. (Cited on page 164.)
- [287] J R Davis, C Y Huang, J Zanet, S Harrison, E Rosten, S Cox, D Y Soong, G A Dunn, and B M Stramer. Emergence of embryonic pattern through contact inhibition of locomotion. *Development*, 139(24):4555–4560, November 2012. (Cited on page 164.)
- [288] DA Kane, M Hammerschmidt, MC Mullins, HM Maischein, M Brand, FJ van Eeden, M Furutani-Seiki, M Granato, P Haffter, and CP Heisenberg. The zebrafish epiboly mutants. *Development*, 123(1):47–55, 1996. (Cited on page 165.)
- [289] Masazumi Tada and Carl-Philipp Heisenberg. Convergent extension: using collective cell migration and cell intercalation to shape embryos. *Development*, 139(21):3897–3904, November 2012. (Cited on page 165.)
- [290] R M Warga and C Nüsslein-Volhard. Origin and development of the zebrafish endoderm. *Development*, 126(4):827–838, February 1999. (Cited on page 165.)
- [291] Ying Gong, Chunhui Mo, and Scott E Fraser. Planar cell polarity signalling controls cell division orientation during zebrafish gastrulation. *Nature*, 430(7000):689–693, August 2004. (Cited on pages 166 and 287.)
- [292] Marion Segalen and Yohanns Bellaïche. Cell division orientation and planar cell polarity pathways. *Seminars in Cell and Developmental Biology*, pages 1–6, May 2009. (Cited on page 166.)
- [293] Zhirong Z Bao, John I JI Murray, Thomas T Boyle, Siew Loon SL Ooi, Matthew J MJ Sandel, and Robert H RH Waterston. Automated cell lineage tracing in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, 103(8):2707–2712, February 2006. (Cited on pages 167 and 295.)
- [294] Willy Supatto, Amy McMahon, Scott E Fraser, and Angelike Stathopoulos. Quantitative imaging of collective cell migration during *Drosophila* gastrulation: multiphoton microscopy and computational analysis. *Nature Protocols*, 4(10):1397–1412, 2009. (Cited on page 167.)
- [295] K Moriyoshi, L J Richards, C Akazawa, D D O’Leary, and S Nakanishi. Labeling neural cells using adenoviral gene transfer of membrane-targeted GFP. *Neuron*, 16(2):255–260, February 1996. (Cited on page 169.)

- [296] Nathan C Shaner, Paul A Steinbach, and Roger Y Tsien. A guide to choosing fluorescent proteins. *Nature Methods*, 2(12):905–909, December 2005. (Cited on page 169.)
- [297] M Westerfield. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish*. University of Oregon, Eugene, January 2000. (Cited on page 169.)
- [298] B Rizzi, M Campana, C Zanella, C Melani, R Cunderlik, Z Krivá, P Bourgine, K Mikula, N Peyrieras, and A Sarti. 3-D Zebrafish Embryo Image Filtering by Nonlinear Partial Differential Equations. *International Conference of the IEEE Engineering in Medicine and Biology Society. Proceedings*, pages 6251–6254, August 2007. (Cited on page 170.)
- [299] Z Krivá, K Mikula, N Peyriéras, B Rizzi, A Sarti, and O Stasová. 3D early embryogenesis image filtering by nonlinear partial differential equations. *Medical Image Analysis*, 14(4):510–526, August 2010. (Cited on page 170.)
- [300] P Frolkovič, K Mikula, N Peyrieras, and A Sarti. Counting number of cells and cell segmentation using advection-diffusion equations. *Kybernetika*, 43(6):817–829, 2007. (Cited on page 172.)
- [301] O Drblikova, M Komornikova, M Remesikova, P Bourgine, K Mikula, N Peyriéras, and A Sarti. Estimate of the cell number growth rate using PDE methods of image processing and time series analysis. *Journal of electrical engineering-Bratislava*, 58(7):86, 2007. (Cited on page 172.)
- [302] A Sarti, R Malladi, and J A Sethian. Subjective surfaces: a method for completing missing boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12):6258–6263, June 2000. (Cited on pages 172 and 173.)
- [303] C Zanella, B Rizzi, C Melani, M Campana, P Bourgine, K Mikula, N Peyrieras, and A Sarti. Segmentation of Cells from 3-D Confocal Images of Live Zebrafish Embryo. *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 6027–6030, 2007. (Cited on pages 172 and 173.)
- [304] Paul Bourgine, Róbert Čunderlík, Olga Drblíková-Stašová, Karol Mikula, Mariana Rešníková, Nadine Peyriéras, Barbara Rizzi, and Alessandro Sarti. 4D Embryogenesis image analysis using PDE methods of image processing. *Kybernetika*, 46(2):226–259, 2010. (Cited on page 173.)
- [305] S Kirkpatrick, C D Gelatt, and M P Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983. (Cited on page 173.)
- [306] S Pauls, B Geldmacher-Voss, and J A Campos-Ortega. A zebrafish histone variant H2A.F/Z and a transgenic H2A.F/Z:GFP fusion protein for in vivo studies of embryonic development. *Development genes and evolution*, 211(12):603–610, December 2001. (Cited on page 174.)
- [307] J L Rubio-Guivernau, V Gurchenkov, M A Luengo-Oroz, L Duloquin, P Bourgine, A Santos, N Peyrieras, and M J Ledesma-Carbayo. Wavelet-based image fusion in multi-view three-dimensional microscopy. *Bioinformatics*, 28(2):238–245, January 2012. (Cited on page 180.)
- [308] J-P Thirion. Image matching as a diffusion process: an analogy with Maxwell’s demons. *Medical Image Analysis*, 2(3):243–260, September 1998. (Cited on page 188.)
- [309] DA Kane and CB Kimmel. The zebrafish midblastula transition. *Development*, 119(2):447–456, 1993. (Cited on pages 211, 212, 213, 214 and 215.)
- [310] L Zhang, C Kendrick, D Julich, and S A Holley. Cell cycle progression is required for zebrafish somite morphogenesis but not segmentation clock function. *Development*, 135(12):2065–2070, May 2008. (Cited on page 214.)
- [311] R T Marler and J S Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6):369–395, April 2004. (Cited on page 222.)

- [312] Lara Carvalho and Carl-Philipp Heisenberg. The yolk syncytial layer in early zebrafish development. *Trends in cell biology*, 20(10):586–592, October 2010. (Cited on page 229.)
- [313] Sherry G Babb and James A Marrs. E-cadherin regulates cell movements and tissue formation in early zebrafish embryos. *Developmental dynamics : an official publication of the American Association of Anatomists*, 230(2):263–277, June 2004. (Cited on pages 242 and 243.)
- [314] Cindy Fukazawa, Celine Santiago, Keon Min Park, William J Deery, Sol Gomez de la Torre Canny, Christopher K Holterhoff, and Daniel S Wagner. poky/chuk/ikk1 is required for differentiation of the zebrafish embryonic epidermis. *Developmental Biology*, 346(2):272–283, October 2010. (Cited on page 244.)
- [315] R O Karlstrom and D A Kane. A flipbook of zebrafish embryogenesis. *Development*, 123:461, December 1996. (Cited on pages 269 and 270.)
- [316] Y Wei and T Mikawa. Formation of the avian primitive streak from spatially restricted blastoderm: evidence for polarized cell division in the elongating streak. *Development*, 127(1):87–96, January 2000. (Cited on page 287.)
- [317] Juliette Mathieu, Kevin Griffin, Philippe Herbomel, Thomas Dickmeis, Uwe Strähle, David Kimelman, Frédéric M Rosa, and Nadine Peyriéras. Nodal and Fgf pathways interact through a positive regulatory loop and synergize to maintain mesodermal cell populations. *Development*, 131(3):629–641, February 2004. (Cited on page 289.)
- [318] Shunji Jia, Di Wu, Cencan Xing, and Anming Meng. Smad2/3 activities are required for induction and patterning of the neuroectoderm in zebrafish. *Developmental Biology*, 333(2):273–284, September 2009. (Cited on page 289.)
- [319] Henri Atlan. *Entre le cristal et la fumée.* essai sur l’organisation du vivant. Éditions du Seuil, 1986. (Cited on page 293.)
- [320] Ivan J Dmochowski, Jane E Dmochowski, Paola Oliveri, Eric H Davidson, and Scott E Fraser. Quantitative imaging of cis-regulatory reporters in living embryos. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12895–12900, October 2002. (Cited on page 295.)
- [321] KR Gabriel and RR Sokal. A new statistical approach to geographic variation analysis. *Systematic Biology*, 18(3):259, 1969. (Cited on page 297.)
- [322] Elijah Flenner, Francoise Marga, Adrian Neagu, Ioan Kosztin, and Gabor Forgacs. Relating biophysical properties across scales. *Current topics in developmental biology*, 81:461–483, 2008. (Cited on page 297.)
- [323] Jonas Ranft, Markus Basan, Jens Elgeti, Jean-François Joanny, Jacques Prost, and Frank Jülicher. Fluidization of tissues by cell division and apoptosis. *Proceedings of the National Academy of Sciences*, 107(49):20863–20868, December 2010. (Cited on page 297.)
- [324] YS Chu, S Dufour, JP Thiery, E Perez, and F Pincet. Johnson-Kendall-Roberts theory applied to living cells. *Physical Review Letters*, 94(2):28102, 2005. (Cited on page 297.)
- [325] Ray Keller. Shaping the vertebrate body plan by polarized embryonic cell movements. *Science*, 298(5600):1950–1954, December 2002. (Cited on page 298.)
- [326] R Keller, LA Davidson, and DR Shook. How we are shaped: The biomechanics of gastrulation. *Differentiation*, 71(3):171, 2003. (Cited on page 298.)
- [327] Ray Keller. Cell migration during gastrulation. *Current opinion in cell biology*, 17(5):533–541, October 2005. (Cited on page 298.)
- [328] Tania Rozario and Douglas W Desimone. The extracellular matrix in development and morphogenesis: A dynamic view. *Developmental Biology*, October 2009. (Cited on page 299.)

- [329] Andras Czirok, Evan A Zamir, Michael B Filla, Charles D Little, and Brenda J Rongish. Extracellular matrix macroassembly dynamics in early vertebrate embryos. *Current topics in developmental biology*, 73:237–258, 2006. (Cited on page 299.)
- [330] Evan A Zamir, Brenda J Rongish, and Charles D Little. The ECM moves during primitive streak formation—computation of ECM versus cellular motion. *PLoS biology*, 6(10):e247, October 2008. (Cited on page 299.)
- [331] Andrew Latimer and Jason R Jessen. Extracellular matrix assembly and organization during zebrafish gastrulation. *Matrix biology : journal of the International Society for Matrix Biology*, October 2009. (Cited on page 299.)
- [332] M A Luengo-Oroz, L Duloquin, C Castro, T Savy, E Faure, B Lombardot, R Bourgine, N Peyrieras, and A Santos. Can voronoi diagram model cell geometries in early sea-urchin embryogenesis? *IEEE*, pages 504–507, 2008. (Cited on page 300.)
- [333] DT Gillespie. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 81(25):2340–2361, 1977. (Cited on page 300.)
- [334] Nina Fedoroff and Walter Fontana. Genetic networks. Small numbers of big molecules. *Science*, 297(5584):1129–1131, August 2002. (Cited on page 300.)
- [335] M Krieg, Y Arboleda-Estudillo, P-H Puech, J Käfer, F Graner, D J Müller, and C-P Heisenberg. Tensile forces govern germ-layer organization in zebrafish. *Nature Cell Biology*, 10(4):429–436, April 2008. (Cited on page 301.)
- [336] Philippe Marmottant, Abbas Mgharbel, Jos Käfer, Benjamin Audren, Jean-Paul Rieu, Jean-Claude Vial, Boudewijn van der Sanden, Athanasius F M Marée, François Graner, and Hélène Delanoë-Ayari. The role of fluctuations and stress on the effective viscosity of cell aggregates. *Proceedings of the National Academy of Sciences of the United States of America*, 106(41):17271–17275, October 2009. (Cited on page 301.)
- [337] Kevin Kit Parker, Amy Lepre Brock, Cliff Brangwynne, Robert J Mannix, Ning Wang, Emanuele Ostuni, Nicholas A Geisse, Josephine C Adams, George M Whitesides, and Donald E Ingber. Directional control of lamellipodia extension by constraining cell shape and orienting cell tractional forces. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 16(10):1195–1204, August 2002. (Cited on page 302.)
- [338] Christian Franck, Stacey A Maskarinec, David A Tirrell, and Guruswami Ravichandran. Three-Dimensional Traction Force Microscopy: A New Tool for Quantifying Cell-Matrix Interactions. *PLoS ONE*, 6(3):e17833, March 2011. (Cited on page 302.)
- [339] Takashi Iwamatsu. Stages of normal development in the medaka Oryzias latipes. *Mechanisms of Development*, 121(7-8):605–618, July 2004. (Cited on page 303.)
- [340] F Graner, B Dollet, C Raufaste, and P Marmottant. Discrete rearranging disordered patterns, part I: robust statistical tools in two or three dimensions. *The European physical journal E, Soft matter*, 25(4):349–369, April 2008. (Cited on page 303.)
- [341] Matteo Rauzi, Pascale Verant, Thomas Lecuit, and Pierre-François Lenne. Nature and anisotropy of cortical forces orienting Drosophila tissue morphogenesis. *Nature Cell Biology*, 10(12):1401–1410, December 2008. (Cited on page 305.)
- [342] Douglas H Erwin and Eric H Davidson. The last common bilaterian ancestor. *Development*, 129(13):3021–3032, July 2002. (Cited on page 306.)
- [343] A Hejnol and M Q Martindale. Acoel development supports a simple planula-like urbilaterian. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1496):1493–1501, April 2008. (Cited on page 306.)

Resumé

Des comportements cellulaires à la morphogenèse embryonnaire : modélisation mécano-génétique et simulations computationnelles du développement animal précoce

Ce travail présente un modèle théorique de morphogenèse animale, sous la forme d'un système complexe émergeant de nombreux comportements cellulaires. Son implémentation repose sur un système multi-agents fondé sur le couplage mécano-génétique entre les dynamiques de la mécanique cellulaire et de la régulation génétique et moléculaire. Notre objectif est l'intégration des mouvements collectifs de myriades de cellules avec les dynamiques d'expression génétique sous-jacentes aux motifs des champs morphogénétiques. Nous examinons aussi les relations de causalité ascendante reliant les comportements cellulaires locaux aux deformations tissulaires globales. Le comportement mécanique de chaque cellule est associé à leur constitution moléculaire et génétique. Nous nous intéressons particulièrement au phénomène d'intercalation cellulaire induit par des comportements de protrusion active. Différents épisodes morphogénétiques se déroulant au cours des 10 premières heures de développement du poisson zébré sont explorés: la période de clivage, la formation de la couche épithéliale externe, l'épibolie, l'internalisation du mésendoderme et la convergence-extension de l'axe antéro-postérieur. Pour chacun de ces phénomènes, une étude de cas examine le rôle respectif des différents tissus impliqués. Les hypothèses que nous proposons sont discutées au moyen de comparaisons automatisées entre les mesures reconstruites issues des données d'imagerie microscopique et des simulations computationnelles. Les reconstructions présentées incluent les champs de déplacements cellulaires *in toto* en 4 dimensions (3D + temps) de l'embryon de poisson zébré. Le modèle est paramétré à travers une interface inspirée du concept de paysage épigénétique de Waddington, permettant de spécifier les phénomènes d'induction et de d'interaction mécanique à l'échelle des champs morphogénétiques.

From Cell Behavior to Embryonic Morphogenesis: Mechanogenetic Modeling and Computational Simulation of Early Animal Development

We present a theoretical model of animal morphogenesis construed as a self-organized phenomenon emerging from a complex system made of a myriad of individual cell behaviors. It is implemented in an agent-based simulation centered on the mechanic-chemical coupling between cellular and genetic dynamics. The goal is to integrate the collective motion of cells and the dynamics of their gene expression underlying the patterning of morphogenetic fields. We also investigate the causal bottom-up link from local cell behavior to global tissue deformation. Each cell's mechanical behavior is mapped from its molecular and genetic identity. Among these behaviors, we focus particularly on cell intercalation as an active process driving tissue deformation. We operate this model to explore the different morphogenetic episode occurring through the first 10 hours of the zebrafish development: cell segmentation, enveloping layer formation, epiboly, internalization and convergence-extension. For each specific episode, a case study is realized to decipher the respective role of the different tissue involved. Quantitative measures reconstructed from both the simulated and the experimental data are compared to automatically explore the multi-dimensional parameter spaces of our hypotheses and their interpretation. Various state of the art computational reconstruction will be presented, including global 4D (3D + time) displacement fields from *in toto* data of the developing zebrafish embryos. A waddingtonian interactive timeline tool to specify intra and inter tissue induction and mechanical behaviors is also proposed.

Mots clés : Evolutionary potential, morphogenesis, modelling, genetic regulation network, cell behavior ontology, data reconstruction workflow, biomechanics