Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 352 (2005) 113–130
PHYSICA A
www.elsevier.com/locate/physa

A cell-centered approach to developmental biology

Roeland M.H. Merks, James A. Glazier*

Department of Physics, Biocomplexity Institute, Indiana University, Swain Hall West 159, 727 East 3rd Street, Bloomington, IN 47405-7105, USA

Available online 13 January 2005

Abstract

Explaining embryonic development of multicellular organisms requires insight into complex interactions between genetic regulation and physical, generic mechanisms at multiple scales. As more physicists move into developmental biology, we need to be aware of the "cultural" differences between the two fields, whose concepts of "explanations" and "models" traditionally differ: biologists aiming to identify genetic pathways and expression patterns, physicists tending to look for generic underlying principles.

Here we discuss how we can combine such biological and physical approaches into a cell-centered approach to developmental biology. Genetic information can only indirectly influence the morphology and physiology of multicellular organisms. DNA translates into proteins and regulatory RNA sequences, which steer the biophysical properties of cells, their response to signals from neighboring cells, and the production and properties of extracellular matrix (ECM). We argue that in many aspects of biological development, cells' inner workings are irrelevant: what matter are the cell's biophysical properties, the signals it emits and its responses to extracellular signals. Thus we can separate questions about genetic regulation from questions about development. First, we ask what effects a gene network has on cell phenomenology, and how it operates. We then ask through which mechanisms such single-cell phenomenology directs multicellular morphogenesis and physiology. This approach treats the cell as the fundamental module of development.

We discuss how this cell-centered approach—which requires significant input from computational biophysics—can assist and supplement experimental research in developmental

*Corresponding author. Tel.: +1 812 855 3735; fax: +1 812 855 5533.
E-mail addresses: post@roelandmerks.nl (R.M.H. Merks), glazier@indiana.edu (J.A. Glazier).
URL: http://biocomplexity.indiana.edu.

0378-4371/$ - see front matter
doi:10.1016/j.physa.2004.12.028

biology. We review cell-centered approaches, focusing in particular on the Cellular Potts Model (CPM), and present the Tissue Simulation Toolkit which implements the CPM.

PACS: 87.18.Bb; 87.18.Hf; 87.18.La; 89.75.Fb; 89.75.Kd

Keywords: Computational biology; Cellular Potts Model; Cell-oriented modeling; Biological ontologies; Developmental biology

1. Introduction

As physicists attempt to apply physical approaches to elucidating biological problems (*biological physics* as opposed to *biophysics*), they need to be sensitive to the fundamental differences between biological and physical epistemology, which reflect the different histories of the two fields [1]. In particular, the apparently innocuous term *model* can cause instant misunderstanding. In biology, an *experimental model* is a standard organism, e.g. a rat or a fruit fly, which exhibits a particular biological phenomenon of interest. A *theoretical model*, or simply a *model* is a conceptual model, which describes a hypothetical mechanism using text and qualitative diagrams of interactions. Such a model is *successful* if it is compatible with qualitative observations and suggests further experiments that could support or reject its hypotheses. We will refer to such models as *biological models* or *conceptual models*.

Biological models, from those describing the structure and function of DNA to the elaborate diagrams describing metabolic and regulatory pathways or the patterning of tissues during development, have had tremendous success and have led to the current era of genomics and proteomics. However, the success of genomics has caused difficulties, which expose the limitations of classical biological models. The success of genomics and the need to simplify sufficiently to develop conceptual models leads to typical statements of the form "Gene X causes cancer," to which a physicist would object that a gene, which is simply an embodiment of information, cannot *cause* anything directly. High-throughput experimental procedures, including DNA sequencing, microarray analyses, and high-throughput mutation studies, have shown that the number of interacting players in biological processes is so large that "mentally juggling" them in conceptual models has become impractical, especially since behaviors usually arise from complex networks of interactions rather than from the behaviors of individual components. Parallel developments in the physics of complex systems have shown that ensembles consisting of large numbers of interacting components exhibit collective behaviors that we cannot always understand intuitively from the behaviors of the individual components. We call such an approach *systems biology* [2] (see Fig. 1), in contrast to a reductionist, genetic essentialist approach.

Epistemologically, biologists have considered explanations to be satisfactory when they can list the elements of a biological process and their connections.

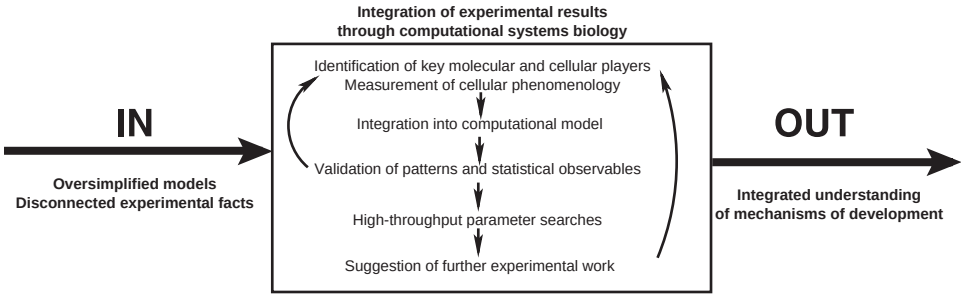

Fig. 1. Schematic description of the systems biology approach.

Metaphorically, the biological model of a car would list all of its parts and their interconnections. A systems-biology approach would include thermodynamics, mechanics, and relevant aspects of chemistry and physics to describe the interactions and functioning of those elements. At the simplest level, the difference in focus is whether we frame questions to answer "what" or "how."

Thus, paradoxically, additional *information* has made the most central biological tool for generating new *knowledge* and *understanding*—the conceptual model—much less helpful. To correctly integrate this wealth of information and to use it to generate new hypotheses, systems biology extends traditional *conceptual models* into *mathematical models*. Because such models are generally too complex for analytic solution, we translate them into sets of algorithms or *computational models*, which we then implement as *simulations*.

These problems are particularly evident when we focus our attention on developmental biology, where a genetic essentialism called *positional information*, which the intellectual school founded by Lewis Wolpert has pioneered and championed, has come to dominate [3]. However useful, the concept of positional information has led to an unfortunate focus on genetic experiment at the expense of other types of measurement. At the same time, even a simple, back-of-the-envelope calculation reveals that development based on the highly specific genetic signaling of positional coding would require much more information than that present in the genome and that the mechanism is not robust to perturbations. Instead, as Turing observed in 1952 [4], development must self-organize under the control of genetic regulation. Cells do not merely sit at a given location and respond passively to external signals, they move, often over long distances, and actively create the signals to which they respond. Thus, the concept of a *prepattern* [5], in which cells migrate towards or respond to a pattern formed by an independent, initiating mechanism has only limited utility. Salazar–Ciudad et al. [6] recently called such prepattern-dependent mechanisms *morphostatic*, as opposed to *morphodynamic* mechanisms in which cell-signaling, migration and growth occur simultaneously, and argued that morphodynamic mechanisms have evolutionary significance because they enable more phenotypic variation for less genotypic variation. Thus, to understand the dynamics of development, and its evolution, we need to return to cell

biology and physiology, which have become somewhat unfashionable in the current genomics rush and in which the experimental technology, while benefiting greatly from advances in gene control, microscopy, microsurgery, labeling and many others areas, has made much less progress than the technology for pure genomics and proteomics.

2. Modeling issues

Computational approaches face their own difficulties, many of which arise from the range of length scales of biological structures and behaviors. Biological phenomena range from those which require quantum-mechanical treatment at the atomic level (10^{-10} m), e.g., the effect of photons on chlorophyll or rhodopsin, through the semi-classical treatment of molecular dynamics (10^{-9} m) and the classical treatment of macromolecules (10^{-8} m), up to agent-based models of populations 10^{3} m).

Multi-scale modeling has not yet had overwhelming success in other disciplines (e.g., weather prediction, materials science and geophysics), and where it has, the methods developed for particular problems have tended not to generalize. The current focus of funding agencies (e.g., the National Science Foundation/National Institutes of Health joint Initiative on Multiscale Modeling) and journals (e.g., the new *Society for Industrial and Applied Mathematics Journal on Multiscale Modeling and Simulation*) on the development of multi-scale technique illustrates its intrinsic importance and difficulty.

Physicists are trained to believe that even very complex phenomena result from the interactions of very few, very simple underlying rules, where the complexity often results from the initial or boundary conditions. A biologist's caricature of a physicist's approach would be to say that the physicist asserts that "development is simply a problem in pattern formation," writes down a set of differential equations and a model of the fertilized egg and tries to deduce a chicken.

Besides the inherent arrogance and doubtful success of such an approach, biology generally lacks the homogeneity which makes physical problems tractable and which permits powerful calculational simplifications like renormalization-group analysis. In addition, evolutionary contingencies mean that universal rules of the type which gives physics its predictive power, need not exist in biology. We cannot assume that simplified models will have the same qualitative behaviors as a more complicated reality, or that observations made in one context will apply to another, even in an apparently similar situation (hence, the widespread feeling among biologists that physical models are "irrelevant"). Thus model validation becomes even more important in biology than in physics and hard-won results may have much more limited ranges of application. On the other hand, we do observe regularities in biology, which give hope that we can indeed generalize, e.g., that most individual cells in a frog are indistinguishable from analogous cells in a human, that certain enzymes and regulatory pathways are identical in yeast and mammals or that cells in all multicellular organisms communicate via secreted products and adhesion.

One specific difficulty is that a simulation can only ever prove sufficiency. It can never prove that a mechanism is correct, nor that a model has included all significant components. Physicists, who believe in Occam's razor, can never directly rebut the complaint "But you have left out ..." Only when simulations can prove their utility by predicting (rather than "post-dicting") experimental observations, will biologists be willing to accept this difference in philosophy.

Thus, biological problems are intrinsically specific, heterogeneous and multi-scale. A computational treatment of a particular problem must begin by choosing an appropriate scale or level of detail, which the inclusion of additional scales can later refine. What is the appropriate first level at which to address development?

Many computational-biology studies of development focus on tissue-level phenomena, modeling tissues as continuous elastic solids or visco-elastic fluids. Others aim to generalize from an understanding of single-cell behaviors and dynamics, building microscopic models of intracellular dynamics (e.g., electrophysiological models or single-cell models of filopodial extension). Some authors (e.g., [7]) argue that coupling many detailed single-cell models can produce models of multicellular phenomena. Molecular and subcellular models like Virtual Cell [8], Silicon Cell[1] or E-cell [9] provide great detail on aspects of subcellular processes. These projects ultimately aim to produce a detailed cell-replica. However worthwhile this aim, such replicas would be just as complex as the biological cells they model and be just as hard to understand. In addition, they can, at best, treat small clusters of a few to tens of cells.

Even if subcellular models replicated aspects of multicellular behavior, could we then say that we understood the behavior? As physicists, we feel that in order to understand a natural phenomenon, we must generate simplifying abstractions organized around common principles of behavior at the appropriate level. Going back to our car metaphor, constructing a functional car replica with LEGO bricks simply by copying one-by-one all the steps in an instruction book does not necessarily teach us *how* a car functions. We obtain insight and understanding by identifying general functional modules like engines, gear-trains, and differentials, observing how they function, how these functional modules interact and distinguishing their universal underlying rules from accidental details such as paint color, choice of metal, or six vs. eight cylinders. Working at too coarse or fine a level of detail makes such an analysis much harder.

The cell provides a natural level of abstraction for mathematical and computational modeling of development. Treating cells phenomenologically immediately reduces the interactions of roughly 10^5-10^6 gene products to 10 or so behaviors: cells can move, divide, die, differentiate, change shape, exert forces, secrete and absorb chemicals and electrical charges, and change their distribution of surface properties.

Is ignoring (to a first approximation) intracellular behaviors legitimate? In a recent book review [10], Meinhardt stated that "The role of the cell as a module of development can hardly be overestimated." Indeed, much of the biology of

[1] http://www.siliconcell.net.

multi-cellular development treats the individual cell's internal properties as a "black box:" for example, the function of an inductive signal does not depend on what happens "behind the scenes," (e.g., on the number of elements in a signal-transduction cascade) as long as the cells respond correctly to the signal. Recently von Dassow and Meir [11] drew a parallel between object-oriented computer programming and modularity in biology. Object-oriented programs consist of modules, the *objects*. These objects have well-defined *interfaces*, through which we address all functionality of the object. In object-oriented programs, *data hiding* is essential: as long as we do not alter the object's response to requests from other objects, we are free to change the object's inner workings. von Dassow and Meir argued that "nature cannot do data hiding" [11], although they conceded that "the closest thing to data hiding [...] is cellular compartmentalization." We argue that nature's solution to data hiding *is* the cell, where signaling molecules (either on the cell surface, or secreted, diffusive signals) and receptors implement the *object interface*. If nature itself uses individual cells as an abstraction, why should we make our biological models more complicated by describing tissue-level structures in terms of subcellular behaviors, which do not affect it directly?

Could we ignore cells and model at the level of tissues? Ignoring cells is dangerous. Macroscopic models, which treat tissues as continuous substances with bulk mechanical properties (e.g., Physiome [12]) reproduce many biological phenomena but fail when structure develops or functions at the cell scale. Although continuum models are computationally efficient for describing non-cellular materials like bone, extracellular matrix (ECM), fluids and diffusing chemicals, many cell-centered models reproduce experimental observations missing from continuum models.

We feel that cell-centered models are essential tools in developmental biology. Cell-centered models use phenomenological models of individual cells to study how the collective behavior of multiple simplified cells drives tissue-level and organism-level processes. The relative simplicity of cell-centered models allows simulation of 10^5-10^6 cells on a single processor, ultimately making whole-organism simulations practical on parallel computers. Fig. 2 shows the hierarchy of scales we include in our models. The questions we can answer with a cell-centered model include: How does the genetic program interact with generic mechanisms to form an organ? What are the relative contributions of local and long-range signaling? What specific factors result in abnormal growth?

Building a cell-centered model requires several steps. First, we infer individual cell behaviors from biological experiments. We can often obtain cell-behavior data from the scientific literature, including the cell's morphology, its response to extracellular and contact-dependent signals, its adhesion to other cell types and the ECM, its chemotactic and haptotactic motility (cell movement in response either to chemical, mechanical or textural gradients in the ECM), the cells into which it can differentiate and the signals which induce differentiation. We may also need to perform additional experiments to obtain precise, quantitative data on cell behavior. Once we have identified how individual cells behave, we can describe the essentials in a conceptual biological model, which we then translate into a mathematical model and implement computationally. This computational description phenomenologically

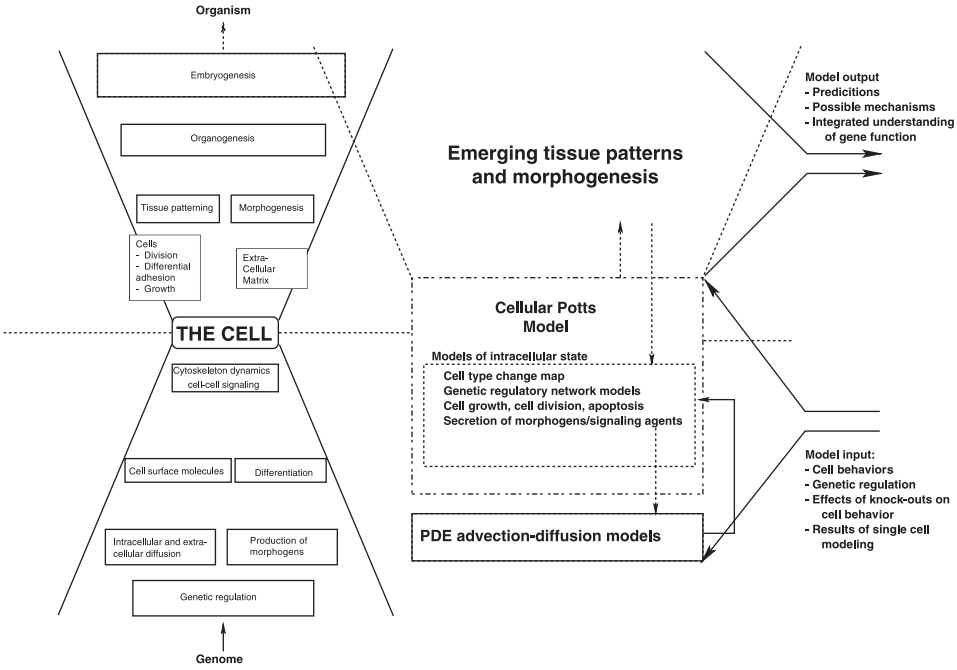

Fig. 2. Cell-centered modeling methodology.

reproduces the individual cells' behaviors. The phenomenological single-cell model is purely descriptive and has no explanatory value per se. It becomes useful when we simulate many single-cell models simultaneously to determine whether the behaviors we included in the single-cell model suffice to explain the tissue-level patterns and physiological functions we find in experiments. If the simulation does not match experimental observations, we return to the experiment to investigate which elements are missing, or which parameter values are inaccurate. High-throughput parameter studies can map out the alternative behaviors of the model, to determine if other parameter combinations better match the experiments. If so, we must either repeat the experimental measurements of sensitive parameters, or determine whether we have missed an essential element of single-cell behavior that a specific parameter choice can mimic. If the model results match experimental observations, we can further test our model by making experimental predictions. What happens to the cell ensemble if we eliminate one of the single-cell behaviors? Can we remove the same element in an experiment, e.g., with a genetic knock-out, and find similar tissue-level ensemble behavior? Thus, by introducing cell-centered simulation into the empirical cycle we can identify the minimal set of single-cell behaviors which suffices to produce certain tissue-level behaviors. We can then ask *which networks of genes* steer this set of single-cell behaviors and *how* they do it. In this way, we can precisely describe and *understand* the role of gene networks in multicellular phenomena,

instead of merely observing that knocking out a gene disrupts a multicellular function.

Several cell-centered computational approaches exist to study morphogenesis in cell aggregates and tissues. These models aim to reconstruct tissue dynamics from the collective behavior of the individual cells. Although some of these approaches keep the position of the cells relative to each other fixed (such as in plant tissues) (see e.g., Refs. [13,14]), most focus on animal tissues, which consist of mobile cells. Such models simulate aggregates of hundreds to tens of thousands of cells, and account for the adhesive forces between cells and between cells and ECM, as well as chemotactic and haptotactic movements. Palsson et al. [15] carry out Newtonian force calculations between individual ellipsoidal cells, to recover the viscoelastic behavior of cell-aggregates and simulate differential-adhesion-driven engulfment of one cell type by another. Another approach is to model aggregations of cells or animals using cellular automata [16] or lattice gases [17–19]). In lattice gases, individual particles that live on a discrete grid represent cells, hopping from one lattice node to the next depending on their discrete velocities. This approach for modeling morphogenesis applies, e.g., to studies of ripple formation in myxomycetes [19], germinal center dynamics [20] and avascular tumor formation [21]. Drasdo et al. [22] introduced a Lagrangian Monte-Carlo method, where attraction, compression and bending energies determine movements of spheroidal cells. Drasdo et al. have used their method to simulate cleavage and gastrulation [23] and avascular tumor growth [24]. Recently, Newman and Grima [25] developed a statistical-mechanics approach to modeling chemotactic cell–cell interactions. Their Langevin-dynamics approach allowed them to study cell ensembles analytically. In the mean-field limit their method recovers the Keller–Segel equations for chemotaxis [26]. Peirce and co-workers used the Netlogo[2] agent-based modeling environment to model cell–cell interactions. In this environment, individual agents move on a square lattice according to a set of preprogrammed rules and the local environment of the agents. Netlogo cannot directly account for cell adhesion, which Peirce et al. explicitly programmed into the agent rules. Peirce et al. applied Netlogo to simulate microvascular remodeling [27] and blastocoel roof thinning during gastrulation of *Xenopus laevis* [28].

The *Cellular Potts Method* (*CPM*) [29] is a convenient and powerful mesoscopic method for modeling cell membrane, cell and tissue dynamics, retaining individual cell identity. Glazier and Graner [29,30] developed the CPM to simulate differential-adhesion-driven cell rearrangement resulting from cell adhesion molecules (CAMs), and quantitatively reproduced cell-sorting experiments.

Additions and improvements to the CPM algorithm include cell growth, cell division, apoptosis and cell differentiation [31], chemotaxis [32], extracellular materials [33] and cell polarity ([34]; Hogeweg, pers. comm.). The CPM is becoming a widely used computational tool in the study of biological morphogenesis, ranging from the full development of the cellular slime mould *Dictyostelium discoideum* [32,33,35], to skeletal formation in the vertebrate limb [36].

[2]http://ccl.northwestern.edu/netlogo.

The remainder of this paper briefly introduces the CPM and illustrates its applications in developmental biology. We conclude with a comparison of the CPM to alternative cell-centered methods.

3. Cellular Potts model

The fundamental choice of the CPM is to represent most cell behaviors in the form of terms within a generalized energy H, which includes the interactions between cells and other cells and the ECM and constraints which determine individual cell behaviors. The genetic regulation of the cell then determines the parameters in these energy terms. We can solve the dynamics of such an energy formalism using a variety of methods, including finite-element, diffuse-boundary, sharp-interface and others. The ordinary CPM we describe below uses a lattice-based Monte-Carlo method with Metropolis dynamics and a generalized Boltzmann weighting function because these are the simplest choices/assumptions. However, we could use off-lattice methods, deterministic dynamics, Kawasaki dynamics and non-Boltzmann weighting functions if they were more computationally efficient or if they better reflected new biological data.

The CPM represents biological cells as patches of lattice sites, $\vec{x}$, with identical indices $\sigma(\vec{x})$, where each index identifies, or "labels" a single biological cell. Connections between neighboring lattice sites of unlike index $\sigma(\vec{x})\neq\sigma(\vec{x}')$ represent membrane bonds, with a characteristic *bond energy* $J_{\tau(\sigma_{\vec{x}}),\tau(\sigma_{\vec{x}'})}$, where the cell types τ (i.e., endothelial, epidermal, etc.) determine the adhesion strength of the interacting cells. An energy penalty increasing with the cell's deviation from a designated target volume A_σ imposes a *volume constraint* on the biological cells. A similar energy penalty on deviations in cell surface area from a target area S_σ constrains the amount of cell membrane. To mimic cytoskeletally driven membrane fluctuations, we randomly choose a lattice site, $\vec{x}$, and attempt to copy its index $\sigma_{\vec{x}}$ into a randomly chosen neighboring lattice site $\vec{x}'$. We reduce the effects of lattice anisotropy by using the 20, first- to fourth-order neighbors on a square lattice. On average, we attempt an update at each lattice site once per Monte-Carlo step (MCS). We calculate how much the energy would change if we performed the copy, and accept the attempt with probability:

$$P(\Delta H) = \{\exp(-(\Delta H + H_0)/T), \Delta H \geqslant -H_0;\ 1, \Delta H < -H_0\}\,, \tag{1}$$

where $H_0 > 0$ is an energy threshold which models viscous dissipation and energy loss during bond breakage and formation [31]. We then define the Hamiltonian as:

$$H = \sum_{\vec{x},\vec{x}'} J_{\tau(\sigma_{\vec{x}}),\tau(\sigma_{\vec{x}'})}(1-\delta_{\sigma_{\vec{x}},\sigma_{\vec{x}'}}) + \lambda\sum_{\sigma}(a_\sigma - A_\sigma)^2 + \lambda'\sum_{\sigma}(s_\sigma - S_\sigma)^2\,, \tag{2}$$

where λ represents resistance to compression, λ' resistance to membrane stretching, a_σ is the current cell volume, s_σ the current cell surface area and the Kronecker delta is $\delta_{x,y} = \{1, x = y; 0, x \neq y\}$. The cells reside in a "medium" which is a generalized CPM cell without a volume constraint and with $\sigma = 0$, $\tau = 0$. We can define the

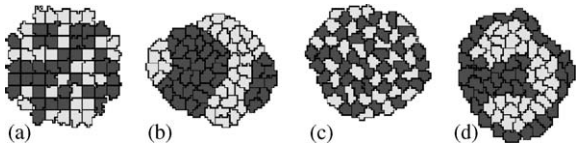

Fig. 3. Simple setup of the Cellular Potts Model (CPM) [29]. (a) Initial condition. Using different bond-strength settings between dark cells (*d*), light cells (*l*) and the surrounding extracellular matrix (*E*), we obtain (b) cell sorting ($\gamma_{ld}>0$, $\gamma_{lE}=\gamma_{dE}>0$), (c) mosaic cell ordering ($0<\gamma_{ld}$, $\gamma_{lE}>0$, $\gamma_{dE}>0$), and (d) engulfment, with the position of high and low adhesivity cells inverted ($\gamma_{ld}>0$, $\gamma_{dE}<\gamma_{lE}$).

surface tensions $\gamma_{\tau_1,\tau_2}=J_{\tau_1,\tau_2}-(J_{\tau_1,\tau_1}+J_{\tau_2,\tau_2})/2$, which enable us to determine whether energetics favors homotypic ($\gamma_{\tau_1,\tau_2}>0$) or heterotypic bonds ($\gamma_{\tau_1,\tau_2}<0$) [29] (see Fig. 3). We usually define a special high *cell-border energy* to prevent cells from adhering artificially to the lattice boundaries. The viscous dissipation H_0 and all terms in the energy, i.e., the bond energies J, and the prefactors to the additional energy terms, such as λ, scale with the temperature T; i.e., if we multiply T by a factor τ, we can multiply H_0 and the energy by the same factor and obtain the same simulation results.

Savill and Hogeweg [32] introduced an elegant method to model *chemotaxis* in the CPM. They used a discretized field C for the concentration of a chemoattractant chemical, defined a set of ordinary differential equations (ODEs) within each cell that describe the cell's secretion and absorption of the chemical, and a set of partial differential equations (PDEs) to calculate the diffusion and decay of the chemical, ignoring advection. They then caused cells to preferentially move up chemoattractant gradients, by adding a term to the energy:

$$H'=H-\sum_{\vec{x},\vec{x}'}\frac{C(\vec{x},t)}{s\ C(\vec{x},t)+1}(1-\delta_{\sigma_{\vec{x},\vec{x}'}}), \tag{3}$$

where the original method sets s, a Michaelis–Menten constant, to 0. Other authors [37,38] have used a similar method to model *haptotaxis*.

Hogeweg [31,39] introduced *cell differentiation*, *cell division* and *apoptosis* to the CPM in a study of the evolution of multicellular development. She used the state of a Boolean network [40]—a simple model of the genetic regulatory network—with inputs from neighboring cells, combined with a simple lock-and-key model of CAMs to determine the bond-energies J. With similar methodologies we are currently introducing more realistic models of genetic regulation into the CPM. Alternative methods for cell differentiation simply use preprogrammed type (τ) changes of the cells [41]. Hogeweg modeled cell division by assigning a new σ to the grid points on one side of the shortest axis of the cell. To implement cell death, she simply set the cell's target volume A to zero, after which the cell vanishes in a few MCSs. Hogeweg implemented cell growth by slowly incrementing the target volume; by consistently doing so when the cell's actual surface area exceeds its target surface area, the cell develops a turgor pressure, which can account for meristimatic growth.

Zajac et al. [34,42] developed a method to account for anisotropic differential adhesion, through the polar distribution of cell-adhesion molecules. They used a shape constraint favoring elongated cells. The bond energy J is highest (weakest adhesion) between two cells touching at their apical and basal surfaces, while it decreases (stronger adhesion) for cells adhering laterally. Zajac's method was slow ($O(N^2)$). However, we have developed an efficient ($O(N)$) calculation of the needed centers of mass and moments of inertia of the cells which runs after each update of the cell boundaries [43].

Most groups using the CPM methodology either developed their own code or hard-coded modifications of previous code. Now, several groups are collaborating to develop an open-source CPM simulation environment that contains the basic functionality CPM models need, and allows easy and transparent extension. Such simulation environments should reduce the overhead in setting up new CPM-based simulations. Just as we do not want to develop from scratch or even tweak the source code of our word processor every time we write a paper, we do not want to redefine the CPM every time we want to set up a cell-oriented simulation. A common environment also allows sharing and archiving of new modules to describe additional biological mechanisms (e.g., an anisotropic adhesion module or detailed models of intracellular genetic regulation). The environment model also allows unified approaches to coupling the CPM with microscopic models like BioSpice[3] and macroscopic models like Physiome. The main current CPM simulation environment is CompuCell3D [36].[4] With CompuCell3D researchers can set up a simulation using the scripting language XML. The environment's architecture is very flexible. Users can add new energy functions, additional constraints, initial conditions and visualization methods by writing plug-ins in C++, analogous to the plug-ins most modern web browsers use. We are currently working to develop a more comprehensive environment, the *Tissue Simulation Toolkit*,[5] combining CompuCell3D with a C++ library for the two-dimensional CPM which we have written in collaboration with Paulien Hogeweg at Utrecht University, The Netherlands. Our aim is to develop a cell-centered simulation environment, which becomes a part of the developmental biologist's workbench, just as bioinformatics and image analysis tools have.

4. The CPM in developmental biology

The collage in Fig. 4 illustrates the wide range of problems life scientists have addressed with the CPM. Hogeweg [31,39] used the CPM to study how evolving genetic regulatory networks interact with generic mechanisms, such as cell sorting, engulfment and intercalation to produce a complicated body plan during development (Fig. 4a). Keşmir and De Boer [44] explained all-or-nothing clonal

[3]https://community.biospice.org/.
[4]http://sourceforge.net/projects/compucell.
[5]http://sourceforge.net/projects/tst.

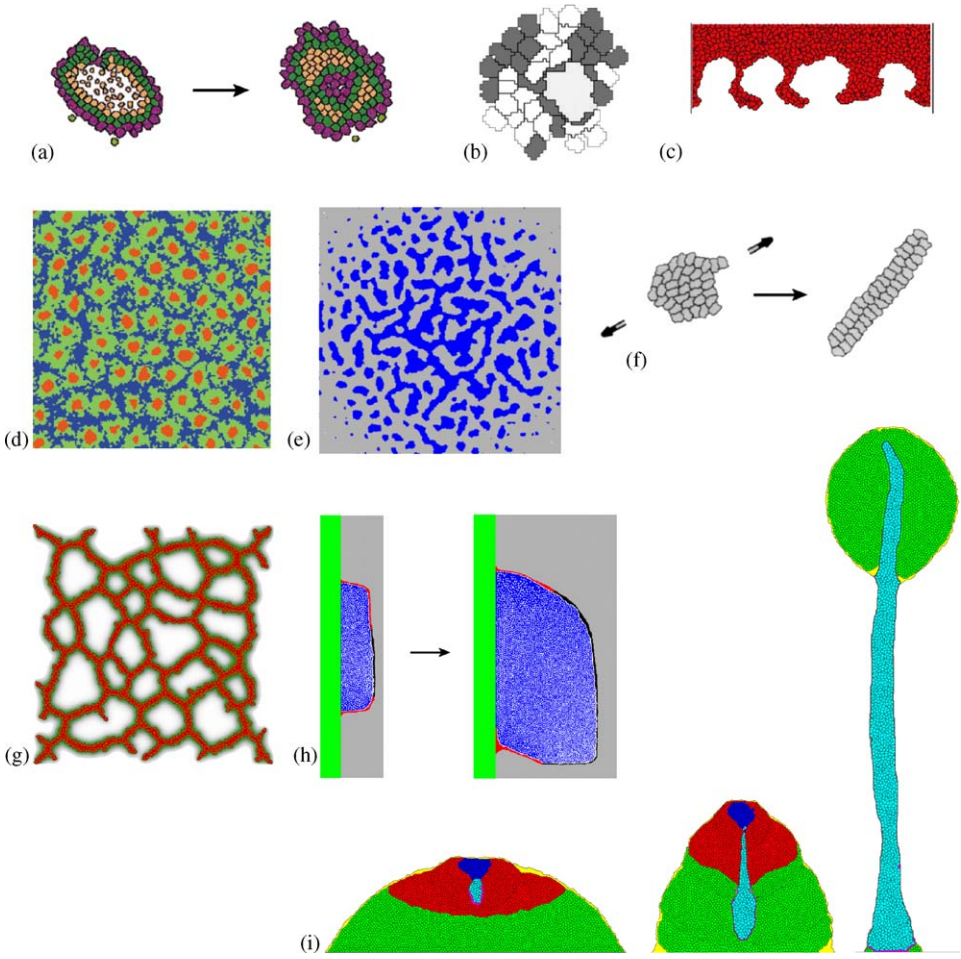

Fig. 4. Cellular Potts modeling in developmental biology. (a) Evolution of developmental mechanisms, we show gastrulation-like mechanism here [31]; (b) clonal selection of β-cells in the germinal center through competition for contact with the (large) antigen-presenting cell [44]; (c) tumor invasion [37]; (d) Notch-Delta mediated stem-cell cluster-size control in the human interfollicular epidermis [41]; (e) mesenchymal condensation through cell–ECM interactions [38]; (f) convergent extension [42]; (g) endothelial cells, secreting VEGF-A, chemotactically aggregate to form a vascular network [43]; (h) limb bud outgrowth, a chemotactic signal from the apical ectodermal ridge (shown in black) causes the asymmetry (Zeng, Thomas and Glazier, unpublished); (i) formation of a fruiting body in *Dictyostelium discoideum* [35].

selection of β-cells during affinity maturation in the germinal center as a competition for adhesive contact to antigen-presenting cells, which they elegantly modeled using the CPM (Fig. 4b). The CPM, like other cell-centered techniques, can provide insight into tumor progression. With the CPM we can straightforwardly evaluate the roles of cell–cell and cell–ECM adhesion in malignant tumor invasion, as Turner and Sherrat [37] and Turner et al. [45] have demonstrated (Fig. 4c).

The CPM is especially useful to study patterning in tissues. Savill and Sherratt [41] combined a subcellular ODE model of membrane-bound Delta-Notch signaling with a discrete-state differentiation model to compare the ability of several hypothetical scenarios to explain clustering of stem cells in the human epidermis (Fig. 4d). Zeng et al. [38] suggested a non-Turing mechanism to explain clustering of mesenchymal cells during cartilage patterning, where cell-secreted fibronectin enhances adhesion between mesenchymal cells and the ECM (Fig. 4e). Extending a previous continuum model [46], Merks et al. [43] developed a cell-centered, chemotaxis-driven model of vasculogenesis (Fig. 4g). Because of the feedback between cell movement, cell adhesion and secretion/diffusion, these cell-level models exhibit qualitative behaviors missing from their continuum PDE limits—an important example of the need for caution in interpreting PDE results.

Fig. 4f, h and i illustrate how we can use the CPM to obtain insight into tissue-level mechanisms of morphogenesis. Zajac et al. [42] modeled convergent extension, a developmental mechanism which, shortly after gastrulation, elongates certain vertebrate embryos (Fig. 4f). They showed that anisotropic intercellular adhesion and cell alignment suffice to produce many aspects of convergent extension. Zeng, Thomas and Glazier (unpublished) showed that chemotaxis toward the apical ectodermal ridge suffices to produce the asymmetric shape of the growing chick limb bud (Fig. 4h). Using a beautiful hybrid CPM–PDE model of *D. discoideum* morphogenesis, Marée and Hogeweg [35] showed that a peristaltic mechanism can explain the formation of the stalk of the fruiting body of *D. discoideum* (Fig. 4i). Pressure waves—caused by the amoebas' chemotactic movement towards excitatory cAMP waves—drive the stalk downwards.

All these studies employed phenomenological models of individual cell behaviors and communication to reproduce and explain experimentally- observed tissue-level behaviors. How can we validate such cell-centered models? First, we must quantitatively compare the simulations with experiments. The CPM provides a wide range of measurements, which we can directly compare with experimental measurements, including tissue patterns, cell positions and velocities, pressure caused by migration or shape changes of cells, cell morphology, surface tensions and morphogen concentrations. We are not yet able, however, to obtain information about the action and reaction forces the cells exert on the medium. We cannot expect cells to take precisely the same migratory paths in simulation and experiment, since they follow different paths in repeated experiments. Instead, we must compare the statistics of cell-migration to the final pattern.

We must be cautious, however, because different biological mechanisms can produce similar patterns. For example, continuum models assuming chemotaxis-driven cell migration [46] and strain-driven cell migration [47] both reproduce aspects of in vitro vascular patterning. We may be able to distinguish between such alternative patterning mechanisms by tracking patterning over time. For example, in a chemotaxis-driven model we expect cells to speed up as they approach each other because the interplay between chemoattractant secretion, diffusion and decay creates exponential chemoattractant gradients around the cells. In a strain-driven model we would instead expect linear dynamics.

Genetic knock-outs are powerful tools for verifying the hypotheses underlying cell-centered models. We can introduce the genetic knock-out's phenomenological effects on individual cell behavior into our cell phenomenology and study how the modified cell behavior alters tissue patterns and mechanisms. We are currently applying such an approach to validate our vasculogenesis model (Fig. 4g). Not only does reproducing the tissue-level effects of knock-outs help validate our computational models, it also helps us unravel the role of particular genetic regulatory networks. Only when we show how and why a gene network is essential in embryonic development, can we say we understand its function.

5. Conclusion and prospects

We have argued how cell-centered simulations, in particular the CPM, can elucidate mechanisms of biological development. We have argued that we can, and indeed must, separate the question of how genetics drives cell behavior from the question of how cell behavior drives morphogenesis. We have previously proposed a similar approach in modeling colonial stony-coral morphogenesis, where we assumed individual coral polyps were the central module of morphogenesis [48,49].

Some authors argue that the energy-minimization approach of the CPM precludes consideration of the forces involved in cell–cell adhesion and cell motion [50]. In the CPM cells move according to effective-energy gradients, $\vec{v} = m\vec{\nabla}E$. Thus, given an effective energy we can calculate the resulting cell motion and the force required to create such motion. In the highly viscous environment of the ECM, force is proportional to velocity, not acceleration, i.e., $\vec{F} = m\vec{v}$, a relation we call the *overdamped force–velocity response*. Thus we can derive effective forces using the relation $\vec{F} = m^2\vec{\nabla}E$. Ongoing work aims to assign physical units to the effective energies. Ultimately we hope to derive CPM parameters, such as those involved in cell adhesion and chemotaxis, from experimentally measured energies involved in the formation and breakage of bonds between cell adhesion molecules, and the effective forces individual filopodia and lamellipodia exert.

Currently, the CPM models the transport of signaling molecules in the ECM with a simple diffusion equation, ignoring advection due to cell and ECM movement. In two dimensions, nutrients or signaling molecules may move through a culture gel over which the cells move (see e.g., Ref. [43]). In three-dimensional simulations we cannot side-step advection. We are now developing a CPM-based advection–diffusion method to fix this unphysical anomaly (D. Dan, K. Chen and J.A. Glazier, to be published). These methods partition the ECM into tiny CPM cells; each of these *fluid cells* contains an attribute describing the current concentration of the advecting and diffusing chemical. A volume constraint ensures that moving biological cells can push the fluid cells. A superimposed diffusion equation describes flux between neighboring fluid cells.

The CPM is just one of a wide range of cell-centered modeling frameworks. Why do we prefer it to these alternative methods? Many cell-centered models describe

individual cells as points or spheres, ignoring that the shape of individual cells can be an essential determining factor in development. In the CPM, a cell's contact energy with surrounding cells and the ECM determines the cell's shape, while the surface area over which two cells adhere determines the forces needed to separate them. Additional constraints can determine specific cell shapes, which may affect tissue-level patterning [42,43].

Another advantage of the CPM is its clear distinction between cell *adhesion* and cell *attraction*. Many cell-centered methods (e.g., Refs. [7,24,50]) approximate cell *adhesion* as cell *attraction*, describing adhesive forces using potentials surrounding the cells (a center-model formalism). However, molecular interactions over very small distances, at least an order of magnitude below the cell scale, cause adhesion. Thus, in a physically correct model of cell adhesion, the cells must touch in order to adhere. The CPM better distinguishes between long-range and short-range attractive and repulsive interactions.

As models and experimental data become more sophisticated, the cell-centered approximation will require extensions to both larger and smaller length-scales. Integrating more microscopic models like BioSpice or Virtual Cell into the CPM is possible in two fashions. They can precalculate or directly control CPM parameters or they can function as components in true hybrid models. Similarly, the CPM can calculate complex materials properties for finite-element continuum models or can interface directly with the finite-element mesh to produce a hybrid model. The key advantage of the CPM is that we need introduce only the minimal additional algorithmic complexity and computation time needed to correspond with the quality of our existing experimental data (hence the suggestive funnel shape we employed in Fig. 2).

Ultimately, we hope that the cell-centered modeling approach will serve developmental biology as a hypothesis-generating and explanatory tool. It will assist developmental biologists to unravel how cell behavior drives tissue-level and organ-level phenomena. Which cell behaviors are essential for observed tissue-level phenomena? Which gene networks are responsible for these essential cell-behaviors? How do gain-of-function and loss-of-function mutations alter individual cell behaviors, and how do these modifications lead to organism-level knock-out phenotypes? Currently, many biologists seem to assume direct causal links between gene action and organism-level phenotypes, and consider that these correlations represent a well-defined gene function. To truly understand gene-*function*, we must put the most essential module of development in the center: the cell.

Acknowledgements

We acknowledge the support from grants NSF-IBN-008365, NASA-NAG2-1619, from Indiana University's Pervasive Technologies Laboratories, Biocomplexity Institute, College of Arts and Sciences and Office of the Vice President for Research and from an IBM Innovation Institute award. We would also like to acknowledge

helpful discussions with Paulien Hogeweg, Wei Zeng, Gilberto Thomas, Stan Marée, Mark Alber, Cornelis Weijer, Stuart Newman, Philip Maini, Santiago Schnell and Mark Chaplain.

References

[1] E.F. Keller, Making Sense of Life, Harvard University Press, Cambrige, MA, 2002.
[2] H. Kitano, Systems biology: a brief overview, Science 295 (2002) 1662–1664.
[3] L. Wolpert, Principles of Development, Oxford University Press, Oxford, 1998.
[4] A.M. Turing, The chemical basis of morphogenesis, Philos. Trans. Roy. Soc. B 237 (1952) 37–72.
[5] J.M.W. Slack, From Egg to Embryo: Determinative Events in Early Development, Cambridge University Press, Cambridge, 1983.
[6] I. Salazar-Ciudad, J. Jernvall, S.A. Newman, Mechanisms of pattern formation in development and evolution, Development 130 (2003) 2027–2037.
[7] T. Krul, J. Kaandorp, J.G. Blom, Modelling developmental regulatory networks, in: P.M.A. Sloot, D. Abramson, A.V. Bogdanov, J.J. Dongarra, A.Y. Zomaya, Y.E. Gorbachev (Eds.), Computational Science—ICCS 2003, Lecture Notes in Computer Science, vol. 2657, Melbourne, Australia and St. Petersburg, Russia, 2003, pp. 688–697.
[8] B.M. Slepchenko, J.C. Shaff, I. Macara, L.M. Loew, Quantitative cell biology with the virtual cell, Trends Cell Biol. 13 (2003) 570–576.
[9] M. Tomita, K. Hashimoto, K. Takahashi, T. Shimizu, Y. Matsuzaki, F. Miyoshi, K. Saito, S. Tanida, K. Yugi, J. Venter, C. Hutchison III, E-CELL: software environment for whole-cell simulation, Bioinformatics 15 (1999) 72–84.
[10] H. Meinhardt, Pathways and building blocks, in: G. Schlosser, G.P. Wagner (Eds.), Review of "Modularity in Development and Evolution", Nature 430 (2004) 970.
[11] G. von Dassow, E. Meir, Exploring modularity with dynamical models of gene networks, in: G. Schlosser, G.P. Wagner (Eds.), Modularity in Development and Evolution, The University of Chicago Press, Chicago, 2004, pp. 244–287.
[12] J.B. Bassingthwaighte, Strategies for the Physiome project, Ann. Biomed. Eng. 28 (2000) 1043–1058.
[13] M. Lantin, An environment for the simulation of biological cells, Ph.D. Thesis, Simon Fraser University, 1999.
[14] D.M. Holloway, M. Lantin, Maintaining apical dominance in the fern gametophyte, Ann. Bot. 89 (2002) 409–417.
[15] E. Palsson, A three-dimensional model of cell movement in multicellular systems, Futur. Gener. Comp. Syst. 17 (2001) 835–852.
[16] M.A. Kiskowski, M.S. Alber, G.L. Thomas, J.A. Glazier, N.B. Bronstein, J. Pu, S.A. Newman, Interplay between activator–inhibitor coupling and cell–matrix adhesion in a cellular automaton model for chondrogenic patterning, Dev. Biol. 271 (2004) 372–387.
[17] H.J. Bussemaker, A. Deutsch, E. Geigant, Mean-field analysis of a dynamical phase transition in a cellular automaton model for collective motion, Phys. Rev. Lett. 78 (1997) 5018–5021.
[18] J.K. Parrish, L. Edelstein-Keshet, Complexity, pattern and evolutionary trade-offs in animal aggregation, Science 284 (1999) 99–101.
[19] U. Börner, A. Deutsch, H. Reichenbach, M. Bär, Rippling patterns in aggregates of myxobacteria arise from cell–cell collisions, Phys. Rev. Lett. 89 (2002) 078101.
[20] M. Meyer-Hermann, A. Deutsch, M. Or-Guil, Recycling probability and dynamical properties of germinal center reactions, J. Theor. Biol. 210 (2001) 265–285.
[21] S. Dormann, A. Deutsch, Modeling of self-organized avascular tumor growth with a hybrid cellular automaton, In Silico Biol. 2 (2002) 1–14.
[22] D. Drasdo, R. Kree, J.S. McCaskill, Monte carlo approach to tissue-cell populations, Phys. Rev. E 52 (1995) 6635–6657.

[23] D. Drasdo, G. Forgacs, Modeling the interplay of generic and genetic mechanisms in cleavage blastulation and gastrulation, Dev. Dyn. 219 (2000) 182–191.
[24] D. Drasdo, S. Hohme, Individual-based approaches to birth and death in vascular tumors, Math. Comput. Model. 37 (2003) 1163–1175.
[25] T.J. Newman, R. Grima, Many-body theory of chemotactic cell–cell interactions, Phys. Rev. E 70 (2004) 051916.
[26] E.F. Keller, L.A. Segel, Model for chemotaxis, J. Theor. Biol. 30 (1971) 225–234.
[27] S. Peirce, E.J. Van Gieson, T.C. Skalak, Multicellular simulation predicts microvascular patterning and in silico tissue assembly, FASEB J. 18 (2004) 731–733.
[28] D. Longo, S.M. Peirce, T.C. Skalak, L. Davidson, M. Marsden, B. Dzamba, De Simone, Multicellular computer simulation of morphogenesis: blastocoel roof thinning and matrix assembly in *Xenopus laevis*, Dev. Biol. 271 (2004) 210–222.
[29] J.A. Glazier, F. Graner, Simulation of the differential adhesion driven rearrangement of biological cells, Phys. Rev. E 47 (1993) 2128–2154.
[30] F. Graner, J.A. Glazier, Simulation of biological cell sorting using a two-dimensional extended Potts model, Phys. Rev. Lett. 69 (1992) 2013–2016.
[31] P. Hogeweg, Evolving mechanisms of morphogenesis: on the interplay between differential adhesion and cell differentiation, J. Theor. Biol. 203 (2000) 317–333.
[32] N.J. Savill, P. Hogeweg, Modelling morphogenesis: from single cells to crawling slugs, J. Theor. Biol. 184 (1997) 229–235.
[33] A.F.M. Marée, P. Hogeweg, Modelling *Dictyostelium discoideum* morphogenesis: the culmination, Bull. Math. Biol. 64 (2002) 327–353.
[34] M. Zajac, G.L. Jones, J.A. Glazier, Model of convergent extension in animal morphogenesis, Phys. Rev. Lett. 85 (2000) 2022–2025.
[35] A.F.M. Marée, P. Hogeweg, How amoeboids self-organize into a fruiting body: multicellular coordination in *Dictyostelium discoideum*, Proc. Natl. Acad. Sci. 98 (2001) 3879–3883.
[36] J.A. Izaguirre, R. Chaturvedi, C. Huang, T. Cickovski, J. Coffland, G. Thomas, G. Forgacs, M. Alber, G. Hentschel, S.A. Newman, J.A. Glazier, CompuCell, a multi-model framework for simulation of morphogenesis, Bioinformatics 20 (2004) 1129–1137.
[37] S. Turner, J.A. Sherratt, Intercellular adhesion and cancer invasion: a discrete simulation using the extended Potts model, J. Theor. Biol. 216 (2002) 85–100.
[38] W. Zeng, G.L. Thomas, J.A. Glazier, Non-Turing stripes and spots: a novel mechanism for biological cell clustering, Physica A 341 (2004) 482–494.
[39] P. Hogeweg, Computing an organism: on the interface between informatic and dynamic processes, Biosystems 64 (2002) 97–109.
[40] S.A. Kauffman, Metabolic stability and epigenesis in randomly connected nets, J. Theor. Biol. 22 (1969) 437–467.
[41] N.J. Savill, J.A. Sherratt, Control of epidermal cell clusters by notch-mediated lateral inhibition, Dev. Biol. 258 (2003) 141–153.
[42] M. Zajac, G.L. Jones, J.A. Glazier, Simulating convergent extension by way of anisotropic differential adhesion, J. Theor. Biol. 222 (2003) 247–259.
[43] R.M.H. Merks, S.A. Newman, J.A. Glazier, Cell-oriented modeling of in vitro capillary development, in: P.M.A. Sloot, B. Chopard, A.G. Hoekstra (Eds.), Cellular Automata. Sixth International Conference on Cellular Automata for Research and Industry, Lecture Notes in Computer Science, vol. 3305, Spinger, Berlin, 2004, pp. 425–434.
[44] C. Keşmir, R.J. De Boer, A spatial model of germinal center reactions: cellular adhesion based sorting of B cells results in efficient affinity maturation, J. Theor. Biol. 222 (2003) 9–22.
[45] S. Turner, J.A. Sherratt, D. Cameron, Tamoxifen treatment failure in cancer and the nonlinear dynamics of TGF-β, J. Theor. Biol. 229 (2004) 101–111.
[46] A. Gamba, D. Ambrosi, A. Coniglio, A. De Candia, S. Di Talia, E. Giraudo, G. Serini, L. Preziosi, F. Bussolino, Percolation morphogenesis and Burgers dynamics in blood vessels formation, Phys. Rev. Lett. 90 (2003) 118101.

[47] J.D. Murray, On the mechanochemical theory of biological pattern formation with application to vasculogenesis, C.R. Biol. 326 (2003) 239–252.
[48] R.M.H. Merks, A.G. Hoekstra, J.A. Kaandorp, P.M.A. Sloot, Polyp oriented modelling of coral growth, J. Theor. Biol. 228 (2004) 559–576.
[49] J.A. Kaandorp, P.M.A. Sloot, R.M.H. Merks, R.P.M. Bak, M.J.A. Vermeij, C. Maier, Morphogenesis of the branching reef coral *Madracis mirabilis*, Proc. Roy. Soc. Lond. B Biol., in press.
[50] E. Palsson, H.G. Othmer, A model for individual and collective cell movement in *Dictyostelium discoideum*, Proc. Natl. Acad. Sci. USA 97 (2000) 10448–10453.