A Robust in Silico Analogue of MDCK Cystogenesis Mimics Growth in Multiple Culture Conditions

Jesse A. Engelberg¹, Anirban Datta², Keith E. Mostov³, and C. Anthony Hunt^{1,3}

¹ Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, California 94143

² Department of Anatomy, University of California, San Francisco, California, 94143

³ Joint Graduate Group in Bioengineering, University of California, San Francisco and Berkeley, California 94143

jesse.engelberg@gmail.com a.hunt@ucsf.edu

Keywords: cell biology, composite agent, cross-model validation, morphogenesis, synthetic modeling, systems biology

Abstract

Madin-Darby canine kidney (MDCK) cells undergoing cystogenesis in vitro is a scientifically useful model of epithelial morphogenesis. The cysts formed in collagen and Matrigel are qualitatively similar, consisting of a single layer of epithelial cells surrounding a hollow lumen. However, differences in key quantitative measures of cyst growth, including cell number and cyst and lumen size, indicate that some cell behaviors are different within the two culture systems. We recently described an agent-oriented analogue of MDCK cystogenesis in Matrigel. It utilized a cellular Potts model and achieved qualitative and quantitative validation targets using empirical parameter tuning. Within this report we highlight steps taken to convert the cellular Potts model framework to one based upon an agent-oriented approach. If measures of cell death are ignored, the only parameters that required adjustment to allow the analogue of cystogenesis in Matrigel to mimic MDCK cystogenesis in collagen were those controlling cell division and polarization. These data indicate that in addition to delayed cell polarization, cell division in collagen is likely slower than in Matrigel. The reported results support the hypothesis that MDCK cells use the same basic operating principles to create cysts when cultured in Matrigel or collagen.

1. INTRODUCTION

In vitro cystogenesis by MDCK cells is a useful model of epithelial morphogenesis and organogenesis. Growth of cysts in culture mimics many behaviors known to occur during epithelial organ development. When MDCK cells are cultured in Matrigel or collagen their behavior follows the same qualitative pattern, with cell division leading to cell polarization, lumen creation, and lumen expansion [1]. However, some features, including cyst size and the onset of lumen formation, are obviously different, as documented in [1].

Available evidence from in vitro experiments does not indicate whether the differences are the result of cells following similar operating principles within different environments, or whether events within collagen cultures are a consequence of operating principles having one or more fundamentally different features.

The development, study, and challenge of agent-oriented analogues of in vitro systems are a new approach to achieving an improved understanding of cell biology. Once properly validated, such analogues provide valuable new insight into mechanisms that may be responsible for referent system behaviors. An in silico MDCK analogue of cell growth within Matrigel cultures (called ISMA-M) was recently reported [2], along with novel quantitative data used for validation. The focus of this report is new in vitro data for MDCK cystogenesis in collagen. The ISMA-M was iteratively refined so that simulation results mimicked most of that data: the resulting analogue is called an ISMA-C. Only parameters influencing cell number and the timing of lumen initiation were altered, yet the new ISMA-C survived strong falsification challenges. Measures of events during simulations provided values for ISMA-C cell number, cyst and lumen size, mean cell area, the ratio of cellular to cyst area, and lumen number percentages that were acceptably similar to those observed in vitro. Within ISMA-C, cyst and lumen size are linked to cell number and onset of lumen formation. We hypothesize that the same is true in collagen cultures. This paper demonstrates that a single in silico analogue, when parameterized differently, can mimic in vitro results generated using different culture systems. It also serves to illustrate the usefulness of the Iterative Refinement Protocol [2-5] and sets the stage for future modeling and simulation efforts.

2. METHODS

Hereafter, to avoid confusing wet-lab with simulated features and behaviors, we use SMALL CAPS when referring to the latter. To generate quantitative data for ISMA-C validation, MDCK cells were grown suspended within collagen (2)

mg/ml) for ten days. Daily samples were fixed and imaged using confocal microscopy, as done previously for Matrigel cultures [2]. Although a higher cell density was used for collagen relative to Matrigel cultures, the frequency of cell clustering was lower. That was because clusters and cysts in Matrigel form on top of a layer of 100% Matrigel as opposed to being fully suspended, as was the case in collagen.

The detailed descriptions in Engelberg et al. [2] for the ISMA system, parameters, and computational methods apply here. A considerably abridged description follows. ISMA-C was constructed so that components and mechanisms map to in vitro counterparts. The system included CELLS, LUMINAL space, and EXTRACELLULAR MATRIX. Components existed on a 2D hexagonal grid in which CELLS and LUMEN occupied multiple grid locations with the remaining space classified as MATRIX (Fig. 1). ISMA-C was implemented using the CompuCell3D [3] cellular Potts model framework and custom code (as described below). Individual CELLS could expand, DIVIDE, change shape, create LU-MENS, move, or change state depending on their internal variables and external environment (Fig. 2). CELLS applied operating principles from specifications developed from a set of targeted in vitro MDCK cystogenesis attributes. Some of these operating principles, such as CELL DIVISION, were modeled on the mechanisms observed within in vitro culture.

An Iterative Refinement Protocol [4-7] was used to adjust CELL behaviors. The IR Protocol is as follows: select a set of targeted attributes that describe behaviors of the referent in vitro system; specify and then implement an analogue to mimic these targeted attributes; falsify the analogue using Similarity Measures (SMs) based on in vitro data specific to the selected targeted attributes; modify the analogue code and/or parameters until it survives falsification. When this is complete, the analogue is validated for that set of targeted attributes. To increase the behavior space captured by the ana-

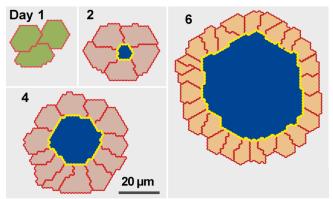


Figure 1. In silico MDCK analogue CYST cross sections. Note that a regular hexagon in hexagonal space maps to a circle in continuous space. Images are from a single simulation run using parameter settings for growth in simulated Matrigel. Growth in simulated collagen produces similar qualitative results. Cells are UNPOLARIZED (green), POLARIZED (gray) or stabilized (orange). Cell-cell and Cell-matrix borders are red; Cell-lumen borders are yellow; Lumens are blue.

logue the IR Process is continued; expand the list of targeted attributes so that the current analogue is falsified by SMs based on in vitro data specific to the new targeted attribute; refine the code and/or parameters until the analogue survives falsification. The new collagen data provided a strong challenge to ISMA-M, which was falsified when using the parameter settings in [2]. To obtain an ISMA-C that would survive falsification, ISMA-M parameter settings were iteratively modified until a satisfactory match to in vitro data was achieved.

In order to mimic the observed smaller cell numbers during the first four days, *clusterProb* was reduced to 0.7 from 0.8 and *cellCycle* was increased from 42 to 100 (the latter maps to 50 hours). Following that refinement the ISMA-C achieved some similarity targets, but LUMENS formed too early. The value of *polarDelay* was adjusted until LUMEN formation in silico matched that in vitro. Surprisingly, no additional changes to parameter settings were required. The underlying code was not modified to produce these results.

The cellular Potts model is a modified cellular automata containing one or more CELLS, each occupying multiple locations on a grid, allowing for the simulation of CELLS with varied size and shape [8, 9]. During a simulation cycle, individual locations on the borders of a CELL can become contained to a neighboring CELL according to probabilistic rules. One rule is that each CELL has an ideal size (number of grid locations), and any grid location change that brings the CELL closer to that ideal is favorable. The CompuCell3D framework implements the CPM, and allows users to simulate biological processes, automating things such as CELL size and

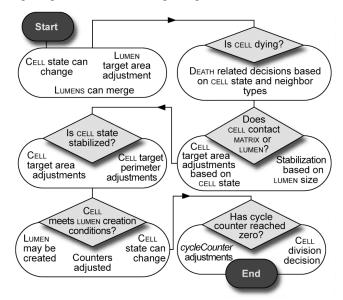


Figure 2. ISMA logic and decision control flow. CELLS step through five logic modules during a simulation cycle, executing actions according to their current state and external environment. Actions include the adjustment of LUMEN target area, CELL DEATH, adjustment of CELL area and state, creation of new LUMENS, and the creation of new CELLS.

Biological terminology

Collagen. A cell culture medium consisting of the collagen I extracellular matrix protein.

Cystogenesis. The process by which cells divide and organize into a hollow cyst consisting of a single (usually) layer of cells. **Epithelial cells.** Cells that make up epithelial organs (such as the skin, lung, kidney), provide a barrier function, and are often polarized, expressing different proteins on either side.

Extracellular matrix (ECM). The protein component surrounding cells that acts as a scaffold for cellular structure.

Lumen Any hollow space, often inside a cyst or duct.

Madin-Darby canine kidney (MDCK) cell. A particular type of dog kidney cell used in cell culture experiments. Capable of surviving in cell culture but lacking most cancerous behaviors.

Matrigel. A complex cell culture medium secreted by Engelbreth-Holm-Swarm (EHC) mouse sarcoma cells containing a number of extracellular matrix proteins.

Morphogenesis. The general process of cellular organization into simple or complex structures.

Polarization. The process or state in which cells express varied proteins in opposite sides, while adjacent sides are in contact with neighboring cells.

Modeling terminology

Biomimetic. Something that imitates biology.

Cellular Potts model (CPM). A cellular automata in which CELLS can occupy multiple locations to model cell size and shape change.

CompuCell3D. A CPM-based framework that automates cell size change and visualization.

ISMA. In Silico MDCK Analogue

ISMA-M. An ISMA parameterized to mimic MDCK cystogenesis within Matrigel culture.

ISMA-C. An ISMA parameterized to mimic MDCK cystogenesis within collagen culture.

MCell. An ISMA object that allows CELLS to access the points contained within them.

Operating principle. A rule or set of rules that translates internal state and environmental cues into observable actions.

Iterative Refinement (IR) process. The process by which the ISMA is refined and improved over time.

Similarity Measure. A numerical measure of the similarity between in vitro and in silico data.

Stabilized. An in silico CELL state in which CELL DIVISION is slowed and other operating principles change. Hypothesized to have an in vitro equivalent.

shape change, CELL ADHESION, CELL DIVISION, and graphical visualization.

The ISMAs used agent-oriented methods. The system contained CELL and LUMEN agents, as well as an agent that executed the underlying cellular Potts model. CELL DIVISION and CELL state change occurred during the execution of CELL agents, while changes to individual grid locations occurred during the execution of the Potts agent. During the latter process, a random set of grid locations was evaluated to see if their index should change from being contained within a specified CELL to being contained within a neighboring CELL or LUMEN. That protocol controlled behaviors such as CELL

size and shape changes, LUMEN expansion, and CELL shrinkage after initiating DEATH. To calculate whether a point would change its index, the Potts agent evaluated the resulting energy change, weighing the changes in CELL area and perimeter, the change in configuration of CELLULAR neighbors, and whether CELLS would be separated or internal rules violated. The latter resulted in an energy penalty. Special rules governing LUMEN expansion prevented CELLS from coming into contact with multiple LUMENS. Each CELL had ideal area and perimeter targets. Changes that created deviations from those targets resulted in higher energies and thus were less likely to occur.

CELLS executed operating principles that dictated when they would change state, DIVIDE, create a LUMEN, or DIE. CELLS had three states: UNPOLARIZED, POLARIZED, and stable. CELL POLARIZATION and DIVISION occurred after a set amount of time elapsed, while LUMEN creation occurred when a POLARIZED CELL contacted other CELLS and the MATRIX, but did not contact a LUMEN. CELLS stabilized when they sensed that their neighboring LUMEN had reached a critical size, controlled by the parameter *stableRatio*. CELL DEATH occurred at random, with CELLS being more likely to DIE when not in contact with the MATRIX.

The stock implementation of CompuCell3D was designed using a grid-based and system-based approach. Individual CELL state change was often initiated after index change events during the Potts execution step. CELLS were not aware of the points contained within them, and there was no way for an individual CELL to execute DIVISION directly. In order to make the process more biomimetic, we developed a custom plug-in that allowed the simulation to be executed from the perspective of individual CELLS. When this plug-in was called, it stepped through each CELL and allowed it to evaluate its internal state and external environment, changing variables such as the target area or polarization state as a result. In addition an MCell object was added to each CELL to create a bidirectional mapping between CELLS and grid points. Each CELL was mapped to an MCell, which contained references to the grid points contained within that CELL. These modifications allowed the analogue to execute using a perspective that was more intuitive, and that mimicked understanding of intracellular biology, where the cell acts as the fundamental functional unit. So doing made modifications significantly easier to implement. For example, CELL DIVI-SION was implemented in an agent-based manner so CELLS had access to information about the points contained within them, thus it was straightforward to randomize or invert the axis of CELL DIVISION.

In order to validate the quantitative results of ISMA-M and ISMA-C, similarity measures were developed for cell number, lumen and cyst size, mean cell area, and the ratio of cellular to cyst area [2]. Similarity Measure 1 (SM1) measured the percentage of in silico simulations at a given day that were within \pm 25% of the mean in vitro value at that day,

thus finding a measure of similarity between in silico and in vitro observations. Based on knowledge of the in vitro system and previous experience designing similarity measures, we judged that if 50% or more simulations fell within that 25% of the in vitro mean (and thus the SM1 value was > 0.5) for 8 out of 10 days, the simulation survived falsification for SM1. The variance of in vitro results can have a significant effect on how stringent the SM1 cutoff should be, so to measure in vitro variance in a way that could be directly compared to SM1, Self-Similarity Measure 1 (SSM1) was developed. SSM1 was exactly the same as SM1, except it measured the similarity between individual in vitro runs and the mean value generated by these runs. Observations with high values of SSM1 had high variance. To directly compare the variance of in silico and in vitro results, rather than simply using the in vitro variance as a guide, Similarity Measure 2 (SM2) was developed. SM2 equaled the absolute value of the difference between the coefficients of variation of the in vitro data and the in silico data at each day. A smaller value of SM2 at a given day indicated that the coefficient of variation was similar for in vitro and in silico data. A simulation survived falsification for SM2 if the value of SM2 was less than 0.25 for 9 out of 10 days for a given measure.

3. EXPERIMENTS AND RESULTS

Cyst growth within collagen was similar to that observed during previous experiments [1]. Suspended cells divided to form clusters, polarized after two to three days, and formed single or multiple lumens by day 3. Lumens and cysts expanded in size thereafter. Within Matrigel cultures, a change in the rate of cell division and cell size was observed at day 6. A similar change was observed within collagen cultures. Growth slowed at roughly day 6, and mean cell size, which had decreased between days 2 and 6, leveled off. After empirical parameter tuning, ISMA-Cs successfully mimicked those qualitative behaviors. Measures of ISMA-C simulations also mimicked quantitative measures of growth within collagen. Prespecified similarity measures (SM1 and SM2) were achieved for CELL number, CYST size, mean CELL area, and ratio of CELLULAR to CYST area, but not for LUMEN size. In order to achieve similarity for CELL number, the parameters clusterProb and cellCycle were altered so that in silico growth matched in vitro growth during days 1-5 (Fig. 3A). ISMA-C simulations produced CELL growth numbers similar to in vitro during days 6-10 without additional tuning. However, CELL stabilization occurred at day 8 instead of at day 6 as observed in vitro.

Cells within collagen polarize and initiate lumens somewhat slower than did their counterparts within Matrigel (Fig. 3B). In order to mimic that behavior, the variable *polarDelay* (which controls the delay before CELLS change to the POLARIZED state) was empirically tuned until mean LUMEN formation time was similar to that observed in vitro. Doing so

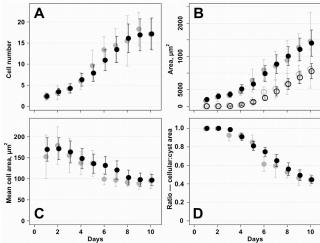


Figure 3. Quantitative values for in vitro and in silico cystogenesis in collagen. Mean values and standard deviations for (A) cell number per cyst, (B) cyst (solid circles) and lumen (open circles) area, (C) mean individual cell area and (D) ratio: cellular to cyst area. Gray: in vitro data taken each day for ten days from 10 cysts. Black: data taken from 50 CYSTS over ten DAYS using specified parameter settings.

required increasing *polarDelay* from 42 (which maps to 21 hours) to 120 simulation cycles (which maps to 60 hours). With that parameter setting, simulations produced mean CYST and LUMEN sizes that were remarkably similar to those observed in vitro (Fig. 3B). It is important to note that variables controlling LUMEN expansion, the effect of CELL stretching on LUMEN expansion, and the mean size of CELLS were not altered, yet ISMA-C results closely mimicked those observed in vitro, indicating that *polarDelay* was the primary driver between these varied simulation results.

Mean cell size within collagen did not differ significantly from that within Matrigel. The good quantitative match between mean ISMA-C and in vitro cell sizes (Fig. 3C) was evidence that the ISMA-C analogue was functioning as intended. Lumen size within in vitro Matrigel and collagen culture differed significantly, especially at later stages of growth. It was thus encouraging to observe that the ratio of CELLULAR to CYST areas during ISMA-C simulations were quite similar to those observed in vitro.

In collagen, the percentage of cysts with multiple lumens was larger than in Matrigel. The percentage of ISMA-C CYSTS with multiple LUMENS was similar to that observed within collagen (Fig. 4A). In addition, the percentage of CYSTS with a single lumen surrounded by a single layer of CELLS (ideal cysts) was also quite similar to in vitro percentages (Fig. 4B). Delaying the formation of ISMA-C LUMENS caused the percentage of CYSTS with multiple LUMENS to increase sufficiently to match in vitro percentages without further parameter adjustments.

Cells within cysts in collagen are more likely to undergo apoptotic cell death than are counterparts in Matrigel [1]. The percentage of ISMA-C CYSTS in which DYING CELLS

were observed (results not shown) did not significantly differ from earlier, published ISMA-M percentages and thus failed to match percentages observed in collagen. Future ISMA-C experiments are needed to explore the consequences of altering parameters that control CELL DEATH in order to obtain frequencies that better match in vitro, while retaining the already achieved attribute similarities.

In order to mimic the differences in cell number and onset of lumen formation observed within collagen cultures, three ISMA-M parameters were altered. The resulting ISMA-C not only mimicked quantitative measures for cell number and lumen initiation in collagen cultures, it also mimicked behaviors for cyst size, lumen size, mean cell size, the ratio of cellular to cyst area, the percentage of cysts with single and multiple lumens, and the percentage of single lumen, single layer cysts.

5. DISCUSSION

In this report we demonstrate use of the Iterative Refinement Protocol. ISMA-M, an analogue developed previously to mimic quantitative measures of MDCK cell growth in Matrigel, was falsified by addition of measures of cystogenesis in collagen to its targeted attribute list. Parameters within the ISMA-M were modified to allow the resulting ISMA-C to survive falsification by most of the new data. Measures taken during ISMA-C simulations quantitatively mimicked corresponding measures within collagen. We believe it is informative to note which parameters had to be altered in order to produce ISMA-C, and which could be left unchanged. To mimic cell numbers, parameters controlling both the initial CELL number and CELL DIVISION rate were altered. With further refinement, we may learn that a higher rate of CELL death will produce later stage numbers that match the lower in vitro cell numbers. However, it is unlikely to be the case during the first four days of growth.

In order to mimic observed values for the onset of lumen formation, the ISMA-M value of *polarDelay* was increased. So doing was in agreement with the hypothesis proposed in [1]: the primary difference between cyst growth in Matrigel and collagen is due to delayed polarization (although our simulation also predicts a decrease in the rate of cellular division). We noted that values of CYST and LUMEN size generated during ISMA-C simulations were remarkably close to values observed within collagen, even though ISMA-C were not modified specifically to achieve that result. That evidence supports our hypothesis that ISMA-C mechanisms have in vitro counterparts.

ISMA-C simulations achieved other validation targets without further modification. We take this as additional evidence that ISMA-C operating principles do have in vitro MDCK counterparts. The percentages of CYSTS with single or multiple LUMENS were remarkably similar to in vitro percentages. Although ISMA-C CELLS POLARIZED and formed

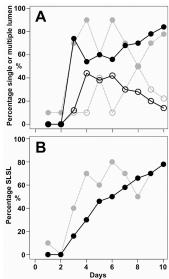


Figure 4. Percentage of cysts with specified types of lumens. (A) Percentage of cysts with single (solid circle) or multiple (open circle) lumens. (B) Percentage of cyst with a single lumen surrounded by a single layer (SLSL) of cells. Gray: in vitro data for 10 cysts taken each day for ten days. Black: in silico data for 50 CYSTS using specified parameter settings. Solid lines: continuous ISMA-C CYST growth. Dotted lines: discrete MDCK cyst growth in collagen.

LUMENS later than in ISMA-M simulations, the number of ISMA-C CELLS within clusters when LUMENS first appeared was similar to that observed in ISMA-M simulations. It may be that the timing of the onset of lumen formation in vitro is primarily a function of cell number.

In order to implement these simulations using agentoriented modeling techniques, it was necessary to reengineer and expand certain aspects of the modeling framework. Cellular Potts models are useful because they can simulate aspects of cell biology like cell size and shape change, but they are not fundamentally agent-oriented. By expanding the CompuCell3D framework to allow direct execution of the CELL agents, we changed the perspective used to create the underlying analogue. Using a perspective consistent with prior simulation efforts allowed concepts and techniques such as the Iterative Refinement Protocol to be fully utilized. It also allowed for the direct mapping from in vitro to in silico operating principles and from in silico observations to in vitro hypotheses. By executing individually and only requiring information about their immediate neighbors, CELLS maintain a degree of autonomy. Increasing the quasi-autonomous nature of agents within simulations will encourage good modeling practices and make mapping from in silico to in vitro more natural and direct.

ISMA-C CELLS stabilize later, at roughly day 8, rather than at day 6 as observed in vitro. This difference indicates that either the ISMA-C mechanisms are a flawed representation of in vitro counterparts, or that a different value of LU-

MEN size is needed. Further, ISMA-C simulations do not mimic the increased rate of cell death observed in vitro. Additional iterative refinement is needed to discover an ISMA-C that does not exhibit those discrepancies. Modification of *stableRatio* is expected to cause CELLS to stabilize earlier. Changes in the parameters influencing CELL DEATH will increase the frequency of CELL DEATH, but that change will impact those features that currently validate. Further successful refinement will expand ISMA-C's phenotype so that it covers a larger portion of the MDCK culture system's in vitro phenotype. The current results support the hypothesis that in achieving cystogenesis, MDCK cells use the same basic operating principles when cultured in Matrigel or collagen.

6. IMPLEMENTATION TOOLS

The model was implemented using CompuCell3D v 3.2.1 (http://compucell3d.org/) and custom code. CompuCell3D is a grid-based cellular Potts model simulation framework. Simulation data was stored in a MySQL v 5.0 (http://www.mysql.com/) database and analyzed using R v 2.10.1 (http://www.r-project.org). Simulations were executed using a Dell Poweredge 1900 server with two 4-core 2.33 GHz 64 bit Intel Xeon processors, 8 GB of RAM and a 450 GB hard drive. The system software was Ubuntu 8.04 LTS (Linux kernel 2.6).

ACKNOWLEDGMENTS

We would like to thank members of the UCSF BioSystems group and the Mostov lab for helpful suggestions and discussion. We gratefully acknowledge research funding provided by the CDH Research Foundation (JAE, CAH), the Alternatives Research and Development Foundation (CAH), and NIH 5R01 DK074398 and 5P01 AI53194 (KEM).

REFERENCES

- [1] Martín-Belmonte, F.; W. Yu; A.E. Rodríguez-Fraticelli; A.J. Ewald; Z. Werb; M.A. Alonso; K. Mostov. 2008. "Cell-Polarity Dynamics Controls the Mechanism of Lumen Formation in Epithelial Morphogenesis." Curr Biol 18:507-513.
- [2] Engelberg, J.A.; A. Datta; K.E. Mostov; C.A. Hunt. 2011. "MDCK Cystogenesis Driven by Cell Stabilization Within Computational Analogues." PLoS Comput Biol (submitted).
- [3] Cickovski T.M.; C. Huang; R. Chaturvedi; T. Glimm; H.G. Hentschel, et al. 2005. "A Framework for Three-Dimensional Simulation of Morphogenesis." *IEEE/ACM Trans Comput Biol Bioinform* 2(4): 273-288.
- [4] Kim, S.H.; J. Debnath; K. Mostov; S. Park; C.A. Hunt. 2009. "A Computational Approach to Resolve Cell Level Contributions to Early Glandular Epithelial Cancer Progression." BMC Syst Biol 3:122.
- [5] Hunt, C.A.; G.E. Ropella; T.N. Lam; J. Tang; S.H. Kim; J.A. Engelberg; S. Sheikh-Bahaei. 2009. "At the Biological Modeling and Simulation Frontier." *Pharm Res* 26(11):2369-2400.
- [6] Tang, J.; C.A. Hunt. 2010. "Identifying the Rules of Engagement Enabling Leukocyte Rolling, Activation, and Adhesion". PLoS Comput Biol 6(2):e1000681.
- [7] Engelberg, J.A.; G.E.P. Ropella; C.A. Hunt. 2008. "Essential Operating Principles for Tumor Spheroid Growth." BMC Sys Biol 2:110.
- [8] Glazier J.A.; A. Balter; N.J. Poplawski. 2007. "Magnetization to morphogenesis: A brief history of the Glazier-Graner-Hogeweg model. In: Rejniak KA, editor. Single-Cell-Based Models in Biology and Medicine. Basel: Birkhäuser. pp. 79-106
- [9] Graner F.; J.A. Glazier. 1992. "Simulation of Biological Cell Sorting Using a Two-Dimensional Extended Potts Model." *Phys Rev Lett* 69(13): 2013-2016.